# REFINEMENT OF OPERATIONAL HEALTH PHYSICS METHODOLOGY USING GEOGRAPHICAL INFORMATION SYSTEMS

By

BRIAN KENT BIRKY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1997

ACKNOWLEDGMENTS

I would like to thank my mentors, Dr. W. Emmett Bolch and Dr. Charles Roessler for guiding me into the rewarding profession of health physics. The patience of the other members of my committee, Dr. Paul Chadik, Dr. David Hintenlang and Dr. William Properzio, throughout this excruciatingly long process is humbly appreciated. Also, I especially thank Mr. Ryan Richards. His knowledge of computer systems and technical assistance with software were crucial to the success of this research. In addition, I would like to acknowledge the contributions of my fellow students over the years: specifically, Kurt Geber, Rich Utrera, Debi Turner, Dave Harrison, Mike Domal, Diane Sinkowski, Susan Stanford, Thabet Tolaymat, Peter LaPuma, George Harder and Tom Huston. The steadfast friendships of Paul Kaiser and Art Schneider are gratefully valued; the former for making me attend virtually every Gator sporting event of the last decade and the latter for his unique ability to stimulate scientific and philosophical curiosities. Finally, I would like to thank my parents, Mr. Edgar Birky and Mrs. Belva Birky, for their love and understanding.

TABLE OF CONTENTS

ACKNOWLEDGMENTS . . . . . ii

Abstract . . . . . v

INTRODUCTION . . . . . 1
Purpose . . . . . 1
Significance . . . . . 2
Literature Review . . . . . 5
An Overview of GIS . . . . . 8
Summary . . . . . 14

PROTOCOL FOR DETERMINATION OF POPULATIONS BY SECTOR-SEGMENT . . . . . 15
Review of Traditional Method . . . . . 15
Data Acquisition and Translation for GIS . . . . . 19
Population Allocation Procedure Using GIS . . . . . 31
Case Study: DOE Pinellas Plant . . . . . 33
Detailed Case Study: St. Lucie . . . . . 37
Summary . . . . . 59

CHAPTER 3
DEVELOPMENT OF POPULATION ESTIMATES AND PROJECTIONS . . . . . 60
Estimate and Projection Concepts . . . . . 60
Methods of Estimation . . . . . 62
Elementary Postcensal Methods of Population Estimation . . . . 70
Population Projections . . . . . 73
Case Study: St. Lucie Estimates and Projections . . . . . 76

CHAPTER 4
DEVELOPMENT OF POPULATION RECONSTRUCTIONS . . . . . 88
Overview . . . . . 88
Methods of Intercensal Population Estimation . . . . . 93
Key Issues with Interpolation Techniques . . . . . 103
Case Study: Pinellas Plant Population Reconstruction . . . . 104
Summary . . . . . 108

CHAPTER 5
ERRORS AND UNCERTAINTIES . . . . . 110
Scope of the Problem . . . . . 110
Dealing with Error . . . . . 120
Error in Digital Representations of Lines . . . . . 127
Polygon Overlay Processing . . . . . 130
Error Estimation for the Sector-Segment Grid . . . . . 130

CHAPTER 6
APPLICATION OF NEW TECHNIQUES TO ENVIRONMENTAL PROBLEMS . . . 136
Overview . . . . . 136

Geographic Criteria for the Siting of Low-Level Waste Disposal Sites . . . . . 137
Environmental Modeling . . . . . 138
Census of Agriculture Data . . . . . 142
Case Study: The Cotter Mill Problem . . . . . 144

CHAPTER 7
APPLICATION OF NEW TECHNIQUES TO EMERGENCY PLANNING . . . . . 155
Overview . . . . . 155
St. Lucie Nuclear Power Plant Evacuation Estimates . . . . . 157
Methodology: Evacuation Times Estimates . . . . . 174

CHAPTER 8
CONCLUSIONS AND RECOMMENDATIONS . . . . . 182
Summary Statement . . . . . 182
Recommended Procedures . . . . . 183

LIST OF REFERENCES . . . . . 188

BIOGRAPHICAL SKETCH . . . . . 198

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

REFINEMENT OF OPERATIONAL HEALTH PHYSICS METHODOLOGY USING GEOGRAPHICAL INFORMATION SYSTEMS

By

Brian Kent Birky

May, 1997

Chairperson: Dr. W. Emmett Bolch
Major Department: Environmental Engineering Sciences

The field of health physics draws upon other specialties for information and technologies relevant to emergency planning and radiation dose assessment. GIS has enjoyed only limited use in health physics to date, but as its applications are demonstrated and the computer hardware and software necessary for its implementation become less expensive, the widespread use by many health physicists is inevitable. This document presents a recommended protocol for the determination of population distributions around nuclear facilities for use as input data for estimating offsite radiation doses due to routine and accidental releases of radioactive materials. Techniques were developed to establish baseline population levels using U.S. Bureau of the Census data and GIS technology, construct current population estimates and future projections using the baseline data coupled with predictive equations and reconstruct past population distributions using current demographic databases. Concurrent to the fulfillment of these objectives is an accounting of uncertainties associated with procedures and parameters. These uncertainties propagate to yield a total uncertainty for the population which in turn is an input parameter for

other models. The significance of this research is that techniques for spatial data analysis are described from a practical perspective. The reader is given the necessary information to choose a GIS platform, obtain population and street mapping data, transform those data into useful formats and interpret results. Proper interpretation of GIS analyses is also important for those professionals who contract with outside agencies for such services. Populations are reported as point estimates with no regard for error in GIS manipulation or input data. The problem stems from the resistance of software vendors to incorporate uncertainty accounting algorithms within their programs and reluctance to release proprietary codes to outside researchers. It is important for health physicists to realize that they can effectively perform many of their own GIS analyses, that the use of census levels that are too aggregated may drastically affect decisions based upon GIS output and that errors and uncertainties are generally ignored in the GIS industry and must be semi-quantitatively estimated.

# CHAPTER 1
# INTRODUCTION

## Purpose

A Geographic Information System (GIS) is defined as:

> a special case of information systems where the database consists of observations on spatially distributed features, activities, or events, which are definable in space as points, lines or areas. A GIS manipulates data about these points, lines, and areas to retreive data for ad hoc queries and analyses. (Deuker, 1975, p. 106)

Many other definitions have been proposed and the ability of a GIS to analyze spatial data is often seen as a key element in its definition. This ability has been frequently used as the characteristic which distinguishes the GIS from systems whose primary objective is map production. The definition chosen to introduce this study exemplifies the analytical approach that made GIS such an attractive tool for this research. GIS technology is quickly breaking the bounds of cartography and geography to become an important analysis tool for many other disciplines in which spatially located data are accumulated. By necessity, the field of health physics draws upon other specialties such as biology, medicine, geology, electronics, computer science and others for information and technologies relevant to emergency planning and radiation dose assessment. GIS has enjoyed only limited use in health physics to date, but as its applications are demonstrated and the computer hardware and software necessary for its implementation become less expensive, the widespread use by many health physicists is inevitable. GIS is a powerful technology that can reasonably be placed

on each professional's desktop. This document presents a recommended protocol for the determination of population distributions around nuclear facilities for use as input data for estimating offsite radiation doses due to routine and accidental releases of radioactive materials. Techniques were developed to accomplish three requirements:

1. establish decennial census population levels using U.S. Bureau of the Census data and GIS technology as a baseline,

2. construct current population estimates and future projections using the baseline data and

3. reconstruct past population distributions using current demographic databases.

Estimation of current and projection of future populations necessarily involves strategies to redirect population growth in regions where saturation, i.e., the attainment of a maximum population density in a user-defined area which is characteristic of the region as a whole, is a problem to adjacent areas where growth can be tolerated. Concurrent to the fulfillment of these objectives is an accounting of uncertainties associated with procedures and parameters. These uncertainties propagate to yield a total uncertainty for the population which in turn is an input parameter for other models.

## Significance

The protocol for determination of population distributions according to a complicated sector-segment grid using a GIS as set forth herein has not been adequately described in previously published literature. The same is true for the projection of future populations using that data structure and the reconstruction of past populations in areas of nuclear weapon production. In addition, a GIS can be used to analyze agricultural production in the area surrounding a nuclear facility so that consumption pathway parameters can be generated that

are site-specific rather than conservative default values. All of these analyses yield point estimates using commercial software. This research addresses this problem in a semi-empirical approach so that default confidence bounds may be applied to results using the same or similar GIS software and similar levels of geographic and census detail. All of the preceding assessments must be performed or contracted for by health physics staff for such vastly important purposes as calculation of collective radiation exposures and radioactivity intakes to the general public, reconstruction of past radiation doses to members of the public at times when dose assessment methodology was not as sophisticated as the present and emergency planning including evacuation modeling in the event of a significant unplanned release of radioactive materials. In short, collective doses based on population distributions assess the public risk and this research offers new techniques to improve those distributions. The collective dose to a population residing in a particular sector-segment is calculated using the series of equations that follow. The first equation calculates the dose to critical organs or tissues of an individual from a particular type and energy of radiation.

$$H_{T,R} = W_R \cdot D_{T,R}$$

where $H_{T,R}$ is the equivalent dose to the organ or tissue (sieverts) from a particular type and energy of radiation

$W_R$ is a dimensionless "radiation weighting factor" which accounts for differences in biological damage caused by different types and energies of radiation

$D_{T,R}$ is the absorbed radiation dose to that tissue or organ in joules per kilogram (gray)

Summing over all types and energies of radiation to which the organ or tissue is exposed gives a total equivalent dose. The next step is to

weight each organ for its own sensitivity to biological endpoints (fatal cancers and other deleterious effects) and sum over all organs and tissues to give a whole body effective dose.

$$E = \sum_T W_T \cdot H_T$$

Where E is the effective dose (sievert) for the whole body

$W_T$ is the tissue weighting factor that accounts for tissue sensitivity

When the radioactivity is inhaled or ingested physical and metabolic factors determine the retention time in the body and the dose calculations are integrated over the subsequent 70 years after intake for a member of the public. This is known as a dose commitment. The collective equivalent and effective dose commitments are the integrals over infinite time of the average individual respective rate quantities due to a specified event, for the critical population group.

$$H_{T,coll} = \int_{t_0}^{\infty} H_{rate,T,avg} \cdot dt$$

Where $H_{T,coll}$ is the collective committed dose equivalent to a particular organ or tissue

$H_{rate,T,avg}$ is the average equivalent dose rate to the tissue for the individuals in the critical population group to be integrated from the time of intake to infinity

The collective effective dose commitment is calculated by summing over all tissues as shown before.

Radiation absorbed dose is commonly measured using integrating thermoluminescent dosimeters (TLDs) in the environs of a facility. The placement and numbers of such monitoring devices are critical for proper

dose assignment and this author recommends that at least one device be placed as close as possible to the centroid of each sector-segment in the population grid as well as known population centers. Since the derived population number directly affects collective dose and these doses are used to assess the operational quality of a plant (and subsequently affect costly insurance premiums), the population should be calculated as accurately as is reasonably achievable. In addition, error bounds should be placed on those populations so that a calculation employing propagation of errors can be performed to provide an uncertainty for equivalent doses produced using them.

## Literature Review

Some of the potential uses of GIS in health physics have been addressed by others in limited detail as follows: data mapping and conjunctive of a global positioning system (GPS) (Runyon, 1994; Sejkora and Most, 1993), visualization (Chen, 1995), emergency response (Mueller, 1995) and waste transportation routing (Sathisan and Chagari, 1994). Runyon states:

> one of the most inportant factors in environmental assessment and characterization is the ability to locate, evaluate, document and relocate sample or data collection sites. Also fundamental to the process is the ability to spatially relate analytical data to those locations. Geographic Information System (GIS) tools for building a spatial data set, for relating data sets and for displaying and querying data are key to the integration process. (Runyon, 1994, p. 9-4)

That paper goes on to describe the power of using both GIS and GPS in tandem and the integration of portable gamma radiation survey equipment. The study site is the area surrounding a non-operational thorium milling facility. Results are presented as radiological isopleths intended to be used for "a multitude of purposes from direction of excavation activities to calculation of dose/risk to

exposed populations" (p. 9-4). This brief treatment does not offer appreciable detail but may be of interest to persons who wish to use GPS with the procedures of Chapter 7 of this dissertation which addresses a more complex environmental assessment. Another recent publication (Sejkora and Most, 1993) also discusses the combination of GIS and GPS with the emphasis on mapping survey locations at environmental thermoluminescent dosimeter placements. The authors refer to vendor GIS data in this statement:

> many of the database maps contain overlays of population and demographic information based on 1990 Census Bureau information, which can be useful in performing dose calculations and emergency response and evacuation planning. (p. 51-52)

What is not stated is that many of these vendor packages of demographic data are restricted to statistical metropolitan areas. They are almost always constructed from the larger census regions, i.e., equal to or larger than the census block group. The census block is the finest level of detail and should always be the starting point for radiological demographic assessments when possible. Later in the paper the authors again mention the use of GIS and population overlays as aids in performing population dose estimates, planning evacuation routes and estimating evacuation times for emergency planning. They further state:

> geographical regions can be defined and delineated within the database, and populations can be estimated in each region. Such regions could be defined using compass sectors, distances, town or city boundaries, major roadways, or combinations of such artificial or existing delineations. (p. 52)

No references that specifically address those applications are given. In contrast, this dissertation provides the detail missing in the existing literature. In a Conference of Radiation Control Program Directors (CRCPD) newsletter Peter Mueller of the Department of Energy's Nevada Field Office reports:

> to support FRMAC (Federal Radiological Monitoring and Assessment Center) deployments and field exercises, geographic databases

> consisting of co-registered "layers" of cultural, radiological, aerial photographs, satellite imagery, and environmental data, are being compiled by EG&G for the areas around all commercial nuclear power plants and Department of Energy facilities. (Mueller, 1995, p. 3)

This is a system in the development phase and appears to use commercial off-the-shelf hardware and software. No publications in support of this system's application regarding procedures have been found. It appears that the system is an aggregation of hardware, software and data; however, actual analysis protocols and data considerations are not described. In a short (4 pages of text) paper, population estimation techniques for routing analysis in support of radioactive waste transport to the proposed repository at Yucca Mountain, Nevada, are investigated (Sathisan and Chagari, 1994). The discussion of the significance of that study points out:

> population characteristics (total population and density) are critical factors in the risk assessment, emergency preparedness and response planning, and ultimately in route designation. (p. 658)

The study compares alternate routes and a series of corridor widths at block, tract and county levels of population aggregation. This extremely limited study purports to examine the impacts of the accuracy or level of disaggregation required for population data. They correctly conclude that aggregation of data over extended lengths of routes results in inaccurate estimation of critical population densities along the routes but erroneously recommend use of tract level data (rather than using block level data) in urban areas. An analysis so important, i.e., the determination of the least hazardous route for radioactive waste transport, must be based on the highest level of detail (the block level). Potential savings in data cost, processing time and storage requirements are trivial in this regard. Therefore, the aggregation level sensitivity analysis is not prudent for their

overall objective. The authors did not describe the data acquisition and translation process. The GIS analysis consisted of buffering various widths around line segments--a simple, standard process for which all such systems are designed thus requiring no particular expertise or explanation.

The preceding publications relating health physics and GIS are few and sparse in detail. There are many texts and journals devoted to GIS. Reviews of that subject and population methodology are broader topics that would not serve this study well. The reader should simply consult the references found in the specific technical discussions presented herein.

## An Overview of GIS

Throughout the last 35 years there has been a very rapid rate of theoretical, technological and organizational development in the GIS field, culminating in a period of intense activity in the last 10 years or so. Many decisions a health physicist must make are dictated or influenced by some aspect of geography. A GIS allows spatial queries that are not possible with other programs such as spreadsheets, statistics packages, or drafting packages. A GIS also links disparate data sets using common fields so that more complex queries are allowed. Linking data sets by identical fields, or exact matching, is a relatively simple process. However, not all data are collected with the same frequency, or the same recording area. For example, environmental data collected weekly can be grouped into monthly categories and linked to a separate database of monthly information. The grouping could also involve small areas aggregated into an exactly matching larger area, e.g., census blocks to counties. Such temporal and areal groupings are

known as hierarchical matching. Note that the result of a hierarchical matching process is an exact match. An exact match is not always a possibility, especially when environmental data and artificial constructs are involved. A common GIS architecture uses layering to view and combine non-matching data sets. The concept is similar to laying transparent maps one over another. With the spatial reference system as the only commonality, layers can be combined to form a new layer with all of their properties. Only then is it possible to perform spatial queries. Combined databases have increased value because they can be used to solve more sophisticated problems. The combination of non-matching layers is known as fuzzy matching.

Once a spatial reference system is in place and spatially referenced data are supplied, some questions that are uniquely answered by a GIS are readily resolved. Inquiries about location, condition, trends, patterns and more complex modeling are particularly suited to this system. To begin an analysis of the area surrounding the Pinellas Plant, for example, the GIS user may enter known latitude and longitude coordinates, or supply the street address so that the site location is found and marked. Places of interest in the vicinity can be located by place name or zip code. Suppose that the user wants to find a new site for a similar facility. The constraints are that the site must be on the coast of Florida with a low population density in a radius of 50 miles. A spatial analysis must be done to find a location where these conditions are satisfied, Once such a location is found, the user might be interested in the rate of growth from 1980 to 1990, i.e., the local trend in population growth. Other factors to consider are background radioactivity and radiation levels and local cancer incidence. When spatially referenced, such data reveal anomalies or patterns that may be of concern. Finally, the GIS can assist in modeling effects of altering

local geographic features for facility construction and modeling theoretical effluent releases and deposition in the environment. These problems are extremely complex and deal with even more complex natural and artificial systems. The use of systems models coupled to GISs creates powerful problem-solving platforms that may represent the most cost-effective way of analyzing and then making decisions concerning environmental issues. However, there exists a danger that the use of GIS technology will grow so much and spread to so many fields that it will forever escape the grasp of the specialists who helped create it.

Data Management

It is not a trivial distinction that the user is part of the GIS. Therefore users need to know the consequences of their selective use of GIS as an analytical tool--not a map-making computer. A database management system (DBMS) is an integral part of all modern geographic information systems. The user's most direct interaction with the basic components of the system is through manipulation of attribute or spatial data through the DBMS. Present DBMS technology is very good for managing tabular data (adding, deleting or modifying records), but even simple mapping involves much more than just the storage and retrieval of data. DBMS technology is not very effective for updating and managing cartographic data, except the attribute data associated with spatial features. This is so because of the topological relationships between map features. When, for example, a land parcel's polygon is adjusted, the relations of surrounding parcels to the first parcel and to each other must also be adjusted; the geometric attributes of all these parcels change.

The Base Map

The Topologically Integrated Geocoding and Referencing (TIGER) File developed by the United States Bureau of the Census can serve as a major component in a geographic information system. The principal cartographic components of the TIGER File are map features (streets, rivers, railroads and so forth) captured in computer-readable form. The Census Bureau accomplished this task by establishing a cooperative agreement with the U.S. Geological Survey to prepare the initial digital base file. Appended to each of these map features (lines) is various geographic information that also serves as a cartographic enhancement; for example, feature names, address ranges, ZIP codes and so forth.

The network of lines formed by these features encloses spaces or areas. The TIGER File contains, appended to these area records, a set of geographic identifiers that also serve as cartographic enhancements; for example, 1990 census block numbers, census tract numbers, city codes, metropolitan statistical area codes and more. The TIGER File, in fact, contains two sets of such codes: one representing the geography at the time of the 1980 census and another representing the geography at the time of the 1990 census. Thus, the first comparison that can be made in a GIS environment is one that determines areal differences from one census to the next.

From a classic GIS standpoint, the geographic area codes provide the mechanism to link the TIGER File's cartographic base to other Census Bureau statistical data sets such as the data tabulations from the 1980 decennial census (recorded on the several 1980 census summary tape files), the 1990 decennial census and the numerous data sets from the several economic, agriculture and governments censuses. The definition of the geographic areas represented in the latter censuses may be

slightly different from what existed in either the 1980 or the 1990 decennial censuses.

The TIGER Files carry latitude and longitude coordinate values for the intersections and end points of the lines comprising the map. These latitude and longitude coordinates permit a GIS to relate the underlying cartographic framework of the TIGER File to data sets from other agencies that are defined by coordinate values; for example, the land use/land cover data sets available from the U.S. Geological Survey, the water resources data and geological data sets also available from the U.S. Geological Survey, the soil data sets from the Department of Agriculture's Soil Conservation Service, the Landsat imagery available from EOSAT Corporation and numerous data sets available from state, county and local units of government that have developed digital files to represent information that interests them.

Attribute Data

The U.S. Bureau of the Census Geography Division provides similar geographic support services for all of the Census Bureau's programs: the decennial census, the economic and agriculture censuses, the intercensal demographic estimates and the current surveys. The scope and complexity of this geographic support task has changed over the decades and the Census Bureau has changed its methodology for conducting censuses and surveys in response.

To meet the geographic support needs of the 1990 Decennial Census, the Census Bureau built a national automated geographic (areal) and cartographic (map) data base called the TIGER file described previously. This file uses a variety of new approaches that deal with the geographic problems associated with the production of over 300,000 individual enumerator assignment maps, appropriate maps for the field supervisory

and office staff, maps to accompany the tabulated results, automated matching of addresses on the mailing lists the Census Bureau uses in many of its data collection activities and preparation of computer files that document the relationship between and among all the geographic tabulation units for the entire United States, its territories and its possessions. The latter is the same set of geographic units that would be seen by looking at a full set of Census Bureau maps spliced together into a single sheet.

Hardware Considerations

GIS applications must access increasingly large data sets, typically several hundred megabytes for a 50-mile sector-segment grid to several gigabytes for regional environmental modeling. The computational activity tends to be I/O dominated, with comparatively little time devoted to mathematical manipulation (Hendley, 1990). Rhind (1988) goes further in explaining the tendency for I/O domination of computational activity by stating that while the type of questions asked of geographical data are conceptually simple, they may involve exhaustive searches of a large database or require multiple prior indexing. This requirement can be illustrated by examining typical GIS operations, such as polygon overlay used in sector-segment population determination, where a single piece of data has to be retrieved and manipulated a number to times in the course of the operation before finally being output (Harding and Hopkins, 1993). GIS algorithms are usually written to take account of the fact that data volume is much larger than system memory. Therefore, these algorithms use local disks for temporary storage, thus compounding the I/O overhead of a GIS application. This tendency towards I/O domination makes disk performance a critical factor in optimizing GIS performance. The

interaction of processor time and I/O is crucial for computational performance. Considering these factors, the GIS application is optimized by increasing the utilization of the processor by minimizing any I/O and increasing the processing capacity of the computer. Since I/O is usually the bottleneck, it must be addressed first, before any benefit from increased processing capacity can be achieved. The bottom line on hardware is to buy a system with the fastest processor available and a large hard drive with fast access and data transfer times. Multiple processor systems are only necessary for extremely large geographical areas, large attribute files and frequent analyses. For most purposes, an off-the-shelf personal computer of high quality will suffice. Select a brand name model since some vendors cut costs by using lower grade components in models that appear to be equivalent, but will perform poorly. Get the fastest processor available coupled with a large (gigabytes in capacity) hard drive with a fast access time and rapid transfer rate. Another necessity is a high speed CD-ROM drive to access Bureau of the Census data.

## Summary

Clearly, the application of GIS techniques to health physics problems is beginning to happen, although the literature discussions to the present have been superficial. The task of performing GIS analyses is falling upon health physicists and their staffs with little guidance. The protocols of this document fill that void. An overview of the GIS concept was presented that serves as a foundation for the more rigorous treatments that follow. The next chapter explains how to acquire geographical and population data, prepare those data for use and analyze those data in a GIS.

CHAPTER 2
PROTOCOL FOR DETERMINATION OF POPULATIONS BY SECTOR-SEGMENT

## Review of Traditional Method

Nuclear facilities routinely record and report all emissions of radioactivity to the environment. Although these emissions are easily quantified via devices such as stack-mounted detectors, the eventual radiation doses to members of the public residing near the facilities are too low to be measured. Modeling is the only methodology to insure compliance with Appendix I, 10CFR50. Mathematical models and the computer programs developed from them require input of the numbers of persons (population group) at a specific location. The population group may be the total number of persons, a particular age grouping, one sex, or a combination of age and sex. These populations are aggregated by polygonal locations defined by directional radii and annular distance from the release site. The population-location array in conjunction with the release data (radionuclides, chemical forms, quantities, rates, dates and times) and pathway-specific data yield collective, individual maximum and average radiation doses to the public. Pathway specific data for aereal emissions include release height, wind speed and stability class. For releases directly to tributaries or potentially to groundwater, the data demands can become more stringent involving parameters such as river flow rate, or depth from ground surface to groundwater, underlying geology, transverse and longitudinal dispersivities, transmissivities, mixing volumes, etc. The resultant

doses carry a degree of uncertainty generated by all the parameters used in their calculation. This will be addressed in Chapter 5. The focus of this chapter is the determination of populations by location for use as a model parameter.

The map overlay methodology is essentially a technique, in environmental planning, to summarize many spatially distributed variables. In "Town and Country Planning Textbook" where Tyrwhitt delineated the scope of a planning survey in the chapter "Surveys for Planning," she also described, almost as an adjunct, how such surveys should be recorded; her discussion may have been the first explicit explanation of the map overlay methodology.

> As far as possible maps should be drawn on transparent paper, so that when completed the maps to the same scale can be "sieved" - i.e. placed one on top of another in turn so that correlations or their absence can be noted. Where relevant "Sieve maps" should be made bringing out the degrees of correlations noted on certain matters. Only when this work is complete can the plan be formed, for, until then, many important factors in determining the lines of the plan will not be brought into focus. (Tyrwhitt, 1950, p. 157)

In the 1960s, Ian McHarg further popularized the overlay methodology by many planning studies based on an overlay process. These included the Staten Island Study, published in "Design with Nature." Later, a study in 1962 by Christopher Alexander and Marvin Mannheim at the Civil Engineering Systems Laboratory of the Massachusetts Institute of Technology entitled "The Use of Diagrams in Highway Route Location: An Experiment" displayed the analytical strength of the methodology in dealing with a complex locational problem. Steinitz et al. (1976, p. 447-448) described the three particularly important elements of this study in relation to development in the map overlay methodology:

> twenty-six different diagrams, each depicting the relative "utility" or desirability of the highway location for one specific requirement of the highway, are considered for analysis. The study recognizes that not all factors can be geographically overlaid at one time, and proposed a formal procedure "tree" for

> the combination and weighting over of overlays.... This study surpasses traditional overlay applications in several ways: in its use of an explicit weighing system for the combination of different components, in the idea of superimposing the different diagrams or maps photographically as separately timed exposures on one print, and in the procedure used to structure the synthesis of the 26 different diagrams.

In their summary remarks, Steinitz et al. reiterated the traditional and fundamental aspects of the map overlay methodology in environmental planning. Although they found, no doubt, the applications at the time of their study had become more complex, diverse and technically sophisticated, Steinitz et al. stated that the use of the overlay approaches differed little between their early development and the applications then.

> ... It is clear that the basic methodology and the underlying logic have changed little. We combine data maps of soil, slope, and other elements in the same manner that Warren Manning probably did in 1912. We commonly hand draw each data element as a complete separate maps through overlays - visually, by redrawing or by photographic reproduction - into some sort of composite representing the analysis. (Steinitz et al., 1976, p. 449)

However, the concept and the approach in map overlay have changed in the 1980s indeed. Ian McHarg had applied the methodology in a manner quite unlike the earlier applications; he had computed various summary statistics, including some cost-benefit measures, for his plans in entirety and for individual components of those plans. These exercises by McHarg offered the obvious opportunity of extending the map overlay methodology beyond its original, purely visually oriented descriptive functions into the truly analytic, prescriptive realms. The drawing and redrawing of zones on the overlay, not just the overlaying of transparent maps; the generation of design alternatives, not just the evaluation of alternatives can be done with a combination of design intuition and analytic thoughts. The preceding description of the evolution of geographical analysis parallels the homologous process for determination of populations around nuclear facilities.

Population and other demographic data should reflect the most accurate methodologies available and should be presented in the most professional manner possible. The importance of accurate estimations of collective radiation doses to non-occupational populations demands the modernization of standard techniques. The traditional method for determination of population distributions can be summarized in the steps below.

1. Transfer census tract and block group boundaries by hand from library maps to USGS quadrangle maps,
2. check county highway maps for more recent revisions,
3. overlay sector-segment lines and arcs by hand,
4. assign percentages of block groups to sector-segments by using ratios of domiciles counted in the block groups,
5. multiply fractions by populations for the block groups making up a sector-segment and sum in a spreadsheet,
6. use published economic and demographic forecasts to construct growth curves for counties,
7. apply growth factors to sector-segments to predict future distributions and
8. plot current and predicted distributions on a radial graph.

For the 1990 census, the Census Bureau developed the TIGER base map as previously described. The TIGER File contains vertices (lines and points) that define items such as roads, voting districts, census regions and their names and address ranges. The average file size for a state is 500 megabytes (7.5 Mb per county). The population data are on compact disks (P.L. 94-171) containing 28 variables such as population by race and age. The computer hardware and software necessary to take advantage of these databases are now commonly available.

Data Acquisition and Translation for GIS

TIGER files are available on CD-ROM directly from the U.S. Bureau of the Census or from federal repositories. The compact disks are labeled with the State(s) they contain. If the user plans to frequently access many locations within the U.S., it may be advantageous to purchase a complete set from the Bureau. The TIGER data supply a geographical framework (albeit planar) useful for visual reference, i.e., the sector-segment grid can be overlain like the traditional hand method over a printed map. The files that must be used for population analyses are the census boundary files which contain no information on roads, waterways, etc.

The GIS software of choice for this investigation is "Atlas GIS" developed by Strategic Mapping, Inc. Geographic census boundary data from the census bureau must be first translated to an acceptable form (BNA file format) and then imported for use in this platform. The first step may be accomplished using "BdryBNA Translator" available from Micro Map & CAD. Use of the BdryBNA Translator is straightforward. Execute the BdryBNA.EXE program and answer a series of prompts. The first prompt asks for the name of the file(s) to be translated including the full path and file name of the TIGER file(s). Use of wildcard characters (* or ?) is permitted for multiple file translation at a single command. Pressing the <ESC> key will exit the program. If standard TIGER file naming conventions are used, the shape coordinate file will automatically be assigned by substitution of the "F42" extension for the "F41" extension of the basic data TIGER file. If this file is not present, the translator will now prompt the user for the shape coordinate file name. The next prompt asks for a choice of one of there projection options that are listed below.

1. None — This outputs the data points in geographic coordinates, i.e., longitude and latitude.

2. Mercator — This outputs data points in the Mercator Projection using a spherical model and Central Meridian of 75.00 West Longitude.

3. UTM — The Universal Transverse Mercator Projection is the projection used by the U.S. Geological Survey for many of their maps and data sets.

The last prompt asks for a choice of boundary options. There are 24 options of which the following listed by code in Table 2-1 are relevant to this study.

Table 2-1 Boundary Options

| | |
|---|---|
| 1 = County | 13 = AIR |
| 2 = Zip | 14 = ANRC |
| 3 = Census Tract | 17 = County 1980 |
| 4 = Block Group | 22 = Census Tract 1980 |
| 5 = Blocks | 23 = Block Group 1980 |
| 12 = AIR FIPS | 24 = Block 1980 |

An alternate command line mode allows multiple file processing utilizing the DOS batch file capability. Run the BdryBNA program with the following parameters:

BdryBNA <Basic FileName> <Shape File Name> <Additional Codes File> <Projection Choice> <Boundary Choice> <Out FileName>

1. Basic File Name: Valid DOS path filename and extension for Basic Data File Type 1.

2. Shape File Name: Valid DOS path filename and extension for Shape Data File Type 2.

3. Additional Codes: Valid DOS path filename and extension for Additional Codes File Type 3.

4. Projection Choice: 1 = None, 2 = Mercator, 3 = UTM

5. Boundary Choice: Codes in Table 2-1.

6. Output File Name: Optional destination filename for BNA output.

For example, a command line for processing the Waynesboro, VA TIGER file to produce the "Waynes.BNA" file of census tract boundaries in UTM projection for Atlas GIS might look like the following line.
BdryBNA TGR51820.F41 TGR51820.F42 TGR51820.F43 2 3 Waynes

The BdryBNA Translator converts basic data and shape coordinates into BNA regions and curves. The BNA feature naming utilizes the codes found in the TIGER Basic Record File. The codes listed in Table 2-2 are a selection of those deemed useful for this study.

Table 2-2 TIGER codes

| Code and Region | Description |
|---|---|
| 1. County: | State FIPS Code + County FIPS Code 51820 |
| 2. ZIP: | Five digit postal ZIP Code 80226 |
| 3. Tract: | State + County + Tract Number + Tract Suffix 518200106.00 |
| 4. Block Group: | State + County + Tract Number + Tract Suffix + Block Group 518200106.002 |
| 5. Block: | State + County + Tract Number + Tract Suffix + Block Number 518200106.00239A |
| 12. AIR FIPS: | State + FIPS American Indian Reservation AIR,TJSA,TDSA,ANVSA 5112345 |
| 13. AIR: | State + 1990 Census American Indian Reservation AIR,TJSA,TDSA,ANVSA 511234 |
| 17. County 80: | State 80 + County 80 51820 |
| 22. Tract 80: | State 80 + County 80 + Census Tract Number 80 + Tract Suffix 518201234.12 |
| 23. Block Group 80: | State 80 + County 80 + Tract Number + Tract Suffix + Block Group 80 518201234121 |
| 24. Block 80: | State 80 + County 80 + Tract 80 + Block 80 518201234.12123 |

The TIGER data sets are compilations from two existing sources, USGS 1:100,000 scale DLG files and 1980 GBF/Dime-files. TIGER files contain both graphic representation information and address ranging. The address ranging is only available for street networks within the GBF/DIME-File coverage of the 1980 Census. Areas outside of these

GBF/DIME areas have been completed with USGS DLG 1:100,000 graphics but do not have address ranging data fields at this time. Distribution coverage is by county or county equivalent. Each county will typically be made up of 12 files shown in Table 2-3.

Table 2-3 TIGER files

| |
|---|
| 1. Basic Data Records . . . Nodes and Geocoding information |
| 2. Shape Coordinates . . . Additional points between nodes |
| 3. Additional Decennial Census Geographic Area Codes |
| 4. Index to alternate Feature Names |
| 5. Feature name list |
| 6. Additional address Range and Zip Code data |
| 7. Landmark Features |
| 8. Area Landmarks |
| A. Additional Polygon Geographic Area Codes |
| I. Area Boundaries |
| P. Polygon Location |
| R. Record Number Range |

BNA is a simple ASCII file format for importing and exporting boundary line data for use in ATLAS GIS (AGIS). Text data in a BNA file describes region, curve, circle and point entities. Once a data resource is translated to this BNA format it can easily be imported for use with ATLAS mapping products using the Import/Export module available from Strategic Mapping. A simple DOS command line to perform the AGIS Import might look like the following line.

```
AGISIE  TGR51820.BNA  Waynes.AGF /Names 1.
```

The BNA format does not limit the number of boundaries in a single file. The BNA format does, however, limit regions and curves to 4,000 vertices. If any line or polygon in the DLG data exceeds this limit it is automatically broken into two or more individual curves. BdryBNA Translator writes the ID number of any boundary which is not closed into

a text file called "LOG.RPT". This file will include the IDs of any boundaries broken into smaller chains as well as boundaries which are incomplete in the original TIGER data. This file is appended with each successive execution of the BdryBNA Translator.

Map makers have long struggled with the problem of portraying the rounded surface of our Earth on flat media such as paper or computer monitors. Many methods of projecting the Earth's curved surface into X-Y coordinate systems have been devised over the years. Each has its advantages but all are compromises which distort the true three dimensional aspect of the Earth to fit the two dimensions of paper or screen.

TIGER graphic data are distributed by the Census Bureau in a standard geographic coordinate system meaning that each point is defined in space by a longitude and latitude. The geographic coordinate system provides the most flexibility for use within Atlas GIS. BdryBNA Translator adds the option of converting this geographic coordinate system to the widely recognized Mercator projection based on a spherical assumption and a central meridian at 75 degrees west longitude. The BdryBNA UTM projection option adds the capability of overlaying common data sets distributed by the U.S. Geological Survey such as Digital Line Graph and 1:24000 scale Digital Elevation Model data. The UTM coordinate system has been used for many published large scale maps. This coordinate system divides the Earth into 60 North to South segments each of 6 degrees longitudinal width. By dividing the earth into smaller segments the distortions of the Transverse Mercator projection are kept to manageable proportions. Translation to other coordinate systems such as a local State plane system are made possible by the use of a coordinate translation program like Tralaine from Mentor Software, Inc.

Population Data

A census is defined as a complete canvass and enumeration of the population and information is collected pertaining to all responding inhabitants of the geographic area under investigation. A census is conducted in this country every 10 years in years ending in zero. Detailed census data are crucial in the production of population estimates and projections. The Bureau of the Census initiated the State Data Center program in 1976 to establish an information network at the state level to help store and disseminate census data. Each state data center in turn establishes its own network of affiliate agencies throughout the state. The State Data Center program has been a successful one that provides local outlets for census data. A modest fee is usually charged for services involving the processing of data tapes, providing printed reports, or providing printed copies of census reports from media such as Microfiche or electronic media.

Population projection methodologies often require data from two successive census periods to establish certain parameters or to calculate population growth rates from one time period to the next. It is therefore useful to note data products for 1990 that are comparable to or incompatible with 1980 census data products.

Apportionment Data

Data for the latest decennial census were collected on census day, April 1, 1990 and then released on gradually with most data products being released between 1991 and 1993. The first major official data product from the 1990 census must be submitted to the president by December 31 of the census year and fulfills the constitutional mandate to conduct a census of population every 10 years for the purpose of apportioning the House of Representatives. A more important mandate for

this study is the delivery by April 1 of the year following the census to each state legislature and the governor of each state the population counts mandated by the provisions of Public Law (P.L.)94-171. The P.L. 94-171 data consist of total population; population by race for white, black, Asian and Pacific Islander, American Indian, Eskimo and Aleut and other, population of total Hispanic origin; and cross-tabulations of data for persons not of Hispanic origin by race. The P.L. 94-171 data are cross-tabulated for all ages and persons 18 years old and over for many geographic areas such as state, county, place, census tract/block numbering area, block group and block. Data for P.L. 94-171 are released on CD-ROM. These data must be matched to census regions as spatial attributes. With the release of the P.L. 94-171 data, the U.S. Bureau of the Census completes its legal obligation with respect to the release of data, but valuable summary data sets are prepared and released as well.

Computer-Accessible Data

One advantage of computer tape files is that the data are directly available in machine-readable form which may expedite the production of population estimates and projections in some cases. Summary tape files (STF) are extremely useful in the production of estimates and projections. Summary tape files contain much more detail than is usually available in conventional printed reports. Four STF products have been designed for the 1990 census, with STF series 1 and 2 containing 100% data and STF series 3 and 4 containing sample data. Each STF contains more than one data file and is labeled STF 1A, STF 1B and so on. STF 3B contains data at the zip code level and may be of more use than other filtered sets.

The four data files of STF 1 include population and housing counts and characteristics. STF 1A contains data for states and state subareas in a hierarchical file structure down to the block group level. The hierarchical file structure allows users to select data at any level of geography included on the tape (now on CD); for example, one might select population counts at the county level, or for the census tracts defining the county, or for the block groups comprising the census tract. STF 1B provides similar data down to the block level. STF 1C and STF 1D provide similar data at levels that are not very useful for the purposes of this study.

The STF 2 series consists of three files that exhibit different file structures, and contains population and housing data similar to the STF 1 series but in greater detail including data by race. STF 2A is structured around the metropolitan statistical area (MSA) level of geography. Data are presented for each state at the census tract/block numbering area (BNA) for MSAs and the non-MSA portion of states. Data are presented also at the census tract/BNA level by county and places of 10,000 or more inhabitants and for whole census tracts/BNAs.

STF 2B contains similar data in an inventory-type file structure with each geographic area presented without linkage to those above or below it, unlike the linkage found in hierarchical file structures. Data are presented for the following geographic areas in STF 2B: states (including summaries for urban and rural), counties, places of 1,000 or more inhabitants and MCDs of 1,000 or more inhabitants in selected states. Data also will be presented for the state portion of American Indian and Alaska Native areas. STF 3C presents data in a hierarchical file structure for the following geographic areas: United States, census regions, census divisions, states (including summaries for urban and rural), counties, places of 10,000 or more inhabitants, MCDs of 10,000

or more inhabitants for selected states, MCDs of fewer than 10,000 inhabitants in New England MSAS, American Indian and Alaska Native areas, metropolitan statistical areas and urbanized areas.

Data on STF 1 and STF 2 represent 100% items asked on all census forms and these data are more likely to be used in the production of population estimates and projections. STF 3 and STF 4 contain data based an a sample of respondents from the 1990 census, but the content is usually of little use in the preparation of estimates and projections with the exception of STF 3B which provides population and housing data at the 5-digit zip code level for each state. A second exception is a special computer tape file pertaining to migration data containing all intrastate county-to-county migration streams and significant interstate county-to-county migration streams.

In some applications, particularly those involving population projections, it is necessary to have data for an individual level of geography at two points in time. Most often one requires data from the two most recent censuses, but population reconstructions for nuclear facilities may extend to the early days of research. The summary tape files for the 1990 census are similar in content to those released for the 1980 census, but are unavailable for previous censuses. There is no guarantee of a one-to-one correspondence between geographic areas in 1980 and 1990 and in some cases one may count on significant changes. For example, only a portion of the nation was defined in terms of census blocks in 1980, but the entire nation was blocked for the 1990 census. In addition, smaller areas of geography are subject to changing boundaries over time. Incorporated places such as cities or towns may annex additional territory between the two census periods and census tracts or block numbering areas may be redefined from one census to another. Caution should be exercised when comparing data from two

census periods for what appears to be the same area of geography. For a 1980 to 1990 evaluation the first task is to compare census boundaries as supplied in the TIGER files. Not all data needs, however, may be met through census data and other sources may be used as illustrated in the detailed examples of the St. Lucie Power Plant population distribution, population projections and evacutation times estimates.

Sample Survey Data

The Current Population Survey (CPS) is conducted by the U.S. Bureau of the Census on a monthly basis. Initiated in 1942, it was designed in response to the need for current information on unemployment during the years of the Depression. The current version of the CPS focuses on employment and unemployment, other characteristics of the general labor force and the population of the United States as a whole based on a large national sample size. Of particular interest in the CPS are data from several of the monthly supplements such as quarterly data on housing vacancy rates, information on migration and household composition, data on fertility, and data on school enrollment.

Vital Registration Data

The federal government does not actually collect vital event data directly, but rather sets standards for data quality and then collates and reports data for births and deaths from all states. Vital registration data for the states are collected and reported by the National Center for Health Statistics, with annual data reported in a multi-volume publication, "Vital Statistics of the United States." Of particular interest are Volume I, on natality and Volume II, on mortality. Monthly data are also published in a "Monthly Vital Statistics Report" series. Annual reports include data at the county level, and monthly data are presented at the national and state levels.

Vital registration data on births and deaths are frequently of use in the production of population estimates and projections.

Each state operates its own vital registration system and more detailed data are usually reported by the state than are available in the U.S. summaries. One is also likely to have access to vital event data more quickly when obtaining it directly from the state. Local governmental units may serve as an additional source of vital event data and may be preferred when projecting at small geographic levels (subcounty).

Additional Data Sources

U.S. Bureau of the Census population projections are published as part of "Current Population Reports," Series P-25, on population estimates and projections for the United States and individual states and on special populations. Frequently included in the publication are population projections produced at the state level by members of the Federal State Cooperative for Population Projections (FSCPP). FSCPP members are a good source for locally produced population projections, especially when one requires detail below the state level as in this study.

The P-25 series of the "Current Population Reports" also publishes the results of the Bureau of the Census's population estimation program. The population estimation program is somewhat more ambitious than the population projection program. Annual estimates for the United States, all states and all counties are produced and published by the Bureau of the Census. Subcounty estimates have been published on a biennial basis as well. County-level estimates are produced by the Bureau of the Census as part of its work with the Federal State Cooperative for Population Estimates (FSCPE), which may represent an alternative source

of data. Many FSCPE members produce their own population estimates at the state and substate level. More information about the FSCPE reports and their usefulness is presented in Chapter 3.

Additional publications of the U.S. Bureau of the Census that may provide helpful information in the population estimation and projection process include the "Statistical Abstract of the United States," and the "County and City Data Book." The "Statistical Abstract of the United States" is a good general resource for a variety of data in addition to the general sections on population and vital statistics at the state level. The "County and City Data Book" provides general population and economic data at the county level and for selected cities in the United States. Migration data are the most difficult to obtain in the process of producing population estimates and projections. Estimates of migration are provided by the Immigration and Naturalization Service (INS). Raymondo (1988) has provided estimates of the undocumented population at the state level using the Bureau of the Census estimate of 200,000 annual undocumented immigrants and the estimated distribution of the undocumented population from the 1980 census.

The migration file produced by the Bureau of the Census from Internal Revenue Service data on matched tax returns is a useful source of migration data. Symptomatic indicators of migration may be necessary when working at levels below the state. Some of the more frequently chosen symptomatic variables include driver's license applications, motor vehicle registrations, voter registrations, utility hookups and school enrollments. School enrollment data may be obtained from local school districts or from state-level departments of education. Utility data may be more difficult to obtain, especially for privately owned or incorporated utility companies.

Individuals included in group quarters are those in institutional settings of some type, such as college dormitories, prisons, long-term health care facilities, military bases and related types of group housing. These categories are important to assess in activities like the evacuation planning demonstrated in Chapter 7. College dormitory counts are best obtained directly from the institution as is the population for long-term health care facilities. Prison populations are usually available at the state level from a state department of corrections or related agency. The Directorate of Information of the U.S. Department of Defense is a good source for data on military base populations,and each branch of the military has an information office that might provide information for its particular bases.

The U.S. Bureau of the Census provides a referral service identifying the names of both private and public data sources in its "National Clearinghouse for Census Data Services" address list. There are many private vendors that specialize in providing demographic products including current population estimates and projections.

## Population Allocation Procedure Using GIS

The starting point for each estimate is the 1990 population enumerated by the U.S. Bureau of the Census (these numbers have not been adjusted for possible undercount or overcount). Consequently, this is also the starting point for sector-segment projections. TIGER files must be translated to regions to match the P.L. 94-171 database. The translated boundary files are then imported into the mapping software. The P.L. 94-171 data are subsequently matched to the census block regions as map attributes.

A much more accurate estimate of 1990, present and future population distributions can be provided than possible with data and technologies of previous censuses. The mapping software chosen for this study has an imbedded dBASE III+ format and will calculate populations for any polygon constructed by automatically assigning fractions, called "area weighting," of census blocks to sector-segments. The algorithm is less likely to fractionalize a large populace for a county with a finely divided population, i.e., one containing a large number of blocks. In the 1980 Census, Alachua County, Florida contained seven census tracts. It is clear that uniquely defined boundaries that intersect portions of those seven tracts would present difficulties in allocations within the intersection. For the 1990 Census, Alachua County, Florida contains over 4,500 census blocks. Since Alachua County contains a total population of 181,596 persons, this focuses the location of about every group of 41 people by a quantum leap in resolution. The computer hardware and software system must be designed to handle this level of detail.

Grid Construction and Data Assignment

A sector-segment grid is constructed originating at the latitude-longitude coordinates of a facility with 16 sectors centered on the compass points and radii in 1 to 10 miles, or corresponding metric increments, to a total of typically 50 miles or 80 kilometers. This is done by using polygon buffering algorithms imbedded in the software program. This is not a perfect system in that no two grids are exactly the same. An apparently uniform grid can have variation. Demographic data associated with census blocks or block groups are assigned to sector-segments by area-weighting as previously described.

Case Study: DOE Pinellas Plant

Eight counties are involved within the 50-mile radius of the DOE Pinellas Plant. These are: Hardee, Hernando, Hillsborough, Manatee, Pasco, Pinellas, Polk and Sarasota. Over 64,000 census blocks are included in those counties. Since the total number of blocks is large, an extensive number of blocks are contained or fractionated by even the smallest sector-segments in the grid (See Figure 2-1). A sector-segment grid was constructed originating at the latitude-longitude coordinates of the DOE Pinellas Plant with 16 sectors centered on the compass points and radii of 2, 3, 4, 5, 10, 20, 30, 40 and 50 miles. Figure 2-2 shows the site with the top of the page as north and the 50-mile grid providing the scale.

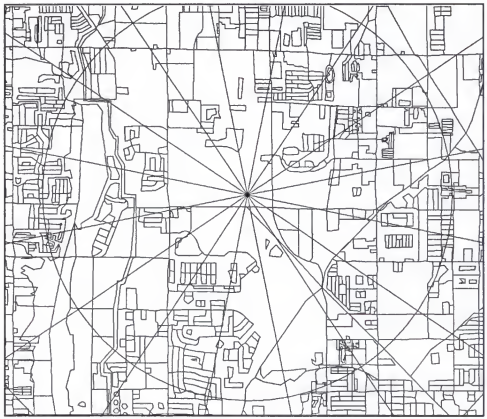

Figure 2-1 Census Blocks Within the One-Mile Grid for the Pinellas Plant

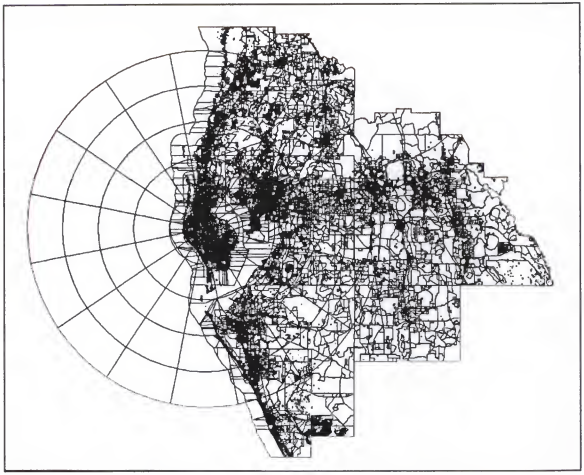

Figure 2-2 Pinellas Plant and the 50-Mile Grid

The allocation of populations to the sector segment grid is accomplished using the following steps.

1. Acquire TIGER CDs for the area of interest and translate the census boundaries according to GIS file format (BNA for Atlas) and map projection (or latitude/longitude pseudoprojection).

2. Acquire P.L. 94-171 data for the same area and extract population data at the summary level in Step 1, e.g., block boundaries and block summary level. The P.L. CDs contain an extraction program to facilitate this procedure.

3. Import translated boundary files into the GIS using an import program such as Import/Export for Atlas.

4. Construct common identity fields between the translated geographic database and the extracted population database so that each census block polygon is linked to its specific population count. It is

sometimes easier (and faster) to use an independent database program to complete this step rather than the DBMS imbedded in the GIS.

5. Use the GIS ring buffering commands to create annular rings at user-specified distances, e.g., 1, 5, 10 miles from a latitude/longitude site coordinate.

6. Use line buffering commands to construct lines passing through the origin point that dissect the rings at user-specified radii, e.g., so that sectors are centered on the sixteen major compass points.

7. Use the "select" command to select each sector-segment to verify individually that each is now an independent buffered region and that no artifact polygons have been created.

8. Perform the GIS analysis "operation" for each sector-segment to provide summary statistics for population (POP100) and other fields of interest, e.g., land area and total area. This step can be automated using a simple script language to write a subroutine.

Tables 2-4 and 2-5 show the 1990 populations by sector-segments. In addition, it might be useful to have cumulative population figures; i.e. 0-2, 0-3, 0-4, 0-5, 0-10, 0-20, 0-30, 0-40 and 0-50 miles for each of the 16 compass directions. The cumulative totals are listed in Tables 2-6 and 2-7. These total are displayed graphically as Figure 2-3.

Table 2-4 Pinellas Plant Populations to 5 Miles

| 1990 POPULATIONS BY SECTOR-SEGMENT | | | | |
|---|---|---|---|---|
| | DISTANCE (MILES) | | | |
| DIRECTION | 0-2 | 2-3 | 3-4 | 4-5 |
| N | 2,421 | 5,829 | 5,564 | 6,763 |
| NNE | 3,806 | 4,228 | 4,394 | 4,121 |
| NE | 2,594 | 1,787 | 4,283 | 1,527 |
| ENE | 1,790 | 599 | 896 | 1,220 |
| E | 2,558 | 1,145 | 831 | 491 |
| ESE | 1,900 | 5,464 | 6,940 | 4,451 |
| SE | 1,818 | 3,523 | 5,711 | 5,581 |
| SSE | 1,597 | 4,457 | 6,592 | 8,947 |
| S | 2,324 | 1,518 | 5,828 | 5,879 |
| SSW | 4,317 | 3,170 | 2,477 | 4,714 |
| SW | 2,288 | 1,591 | 5,751 | 5,237 |
| WSW | 1,575 | 3,417 | 6,663 | 6,661 |
| W | 2,371 | 5,154 | 3,037 | 6,989 |
| WNW | 2,140 | 4,997 | 5,377 | 8,746 |
| NW | 1,553 | 3,515 | 5,692 | 8,266 |
| NNW | 1,013 | 3,446 | 5,945 | 7,894 |

Table 2-5 Pinellas Plant Populations from 5 to 50 Miles

| 1990 POPULATIONS BY SECTOR-SEGMENT | | | | | |
|---|---|---|---|---|---|
| | DISTANCE (MILES) | | | | |
| | 5-10 | 10-20 | 20-30 | 30-40 | 40-50 |
| DIRECTION | | | | | |
| N | 56,502 | 107,728 | 111,008 | 46,903 | 5,324 |
| NNE | 25,146 | 34,356 | 15,540 | 24,095 | 69,288 |
| NE | 5 | 113,351 | 113,027 | 20,064 | 52,096 |
| ENE | 1,417 | 140,450 | 185,621 | 50,789 | 69,815 |
| E | 12,557 | 52,905 | 83,414 | 35,339 | 50,579 |
| ESE | 48,943 | 2,532 | 33,301 | 2,848 | 540 |
| SE | 70,541 | 19,709 | 13,577 | 3,355 | 2,307 |
| SSE | 65,311 | 33,030 | 99,304 | 137,954 | 125,700 |
| S | 30,788 | 7,472 | 12,630 | 2,703 | 0 |
| SSW | 6,530 | 0 | 0 | 0 | 0 |
| SW | 4,692 | 0 | 0 | 0 | 0 |
| WSW | 6,174 | 0 | 0 | 0 | 0 |
| W | 9,052 | 0 | 0 | 0 | 0 |
| WNW | 7,040 | 0 | 0 | 0 | 0 |
| NW | 8,451 | 0 | 0 | 0 | 0 |
| NNW | 41,734 | 2,959 | 0 | 0 | 0 |

Table 2-6 Pinellas Plant Cumulative Populations to 5 Miles

| 1990 POPULATIONS BY SECTOR-SEGMENT CUMULATIVE TOTALS | | | | |
|---|---|---|---|---|
| | DISTANCE (MILES) | | | |
| | 0-2 | 0-3 | 0-4 | 0-5 |
| DIRECTION | | | | |
| N | 2,421 | 8,250 | 13,814 | 20,577 |
| NNE | 3,806 | 8,034 | 12,428 | 16,549 |
| NE | 2,594 | 4,381 | 8,664 | 10,191 |
| ENE | 1,790 | 2,389 | 3,285 | 4,505 |
| E | 2,558 | 3,703 | 4,534 | 5,025 |
| ESE | 1,900 | 7,364 | 14,304 | 18,755 |
| SE | 1,818 | 5,341 | 11,052 | 16,633 |
| SSE | 1,597 | 6,054 | 12,646 | 21,593 |
| S | 2,324 | 3,842 | 9,670 | 15,549 |
| SSW | 4,317 | 7,487 | 9,964 | 14,678 |
| SW | 2,288 | 3,879 | 9,630 | 14,867 |
| WSW | 1,575 | 4,992 | 11,655 | 18,316 |
| W | 2,371 | 7,525 | 10,562 | 17,551 |
| WNW | 2,140 | 7,137 | 12,514 | 21,260 |
| NW | 1,553 | 5,068 | 10,760 | 19,026 |
| NNW | 1,013 | 4,459 | 10,404 | 18,298 |

Table 2-7 Pinellas Plant Cumulative Populations From 5 to 50 Miles

| 1990 POPULATIONS BY SECTOR-SEGMENT CUMULATIVE TOTALS | | | | | |
|---|---|---|---|---|---|
| | DISTANCE (MILES) | | | | |
| DIRECTION | 0-10 | 0-20 | 0-30 | 0-40 | 0-50 |
| N | 77,079 | 184,807 | 295,815 | 342,718 | 348,042 |
| NNE | 41,695 | 76,051 | 91,591 | 115,686 | 184,974 |
| NE | 10,196 | 123,547 | 236,574 | 256,638 | 308,734 |
| ENE | 5,922 | 146,372 | 331,993 | 382,782 | 452,597 |
| E | 17,582 | 70,487 | 153,901 | 189,240 | 239,819 |
| ESE | 67,698 | 70,230 | 103,531 | 106,379 | 106,919 |
| SE | 87,174 | 106,883 | 120,460 | 123,815 | 126,122 |
| SSE | 86,904 | 119,934 | 219,238 | 357,192 | 482,892 |
| S | 46,337 | 53,809 | 66,439 | 69,142 | 69,142 |
| SSW | 21,208 | 21,208 | 21,208 | 21,208 | 21,208 |
| SW | 19,559 | 19,559 | 19,559 | 19,559 | 19,559 |
| WSW | 24,490 | 24,490 | 24,490 | 24,490 | 24,490 |
| W | 26,603 | 26,603 | 26,603 | 26,603 | 26,603 |
| WNW | 28,300 | 28,300 | 28,300 | 28,300 | 28,300 |
| NW | 27,477 | 27,477 | 27,477 | 27,477 | 27,477 |
| NNW | 60,032 | 62,991 | 62,991 | 62,991 | 62,991 |

Detailed Case Study: St. Lucie

This detailed example presents population estimates for the area surrounding the St. Lucie Nuclear Power Plant. It conforms to Nuclear Regulatory Commission (NRC) Regulatory Guide 1.70, Revision 3 requirements. It is provide to demonstrate the complexity demanded in a typical population determination exercise, to illustrate the variety of data that can be spatially analyzed and the display capabilities of the system. The population estimates presented herein represent 1990 resident population levels; and estimated 1992 resident, transient and special populations for the area within 10 miles of the St. Lucie plant.

Specification of Location

Florida Power & Light Company's (FPL) St. Lucie site is located on Hutchinson Island, St. Lucie County, Florida. The coordinates for St.

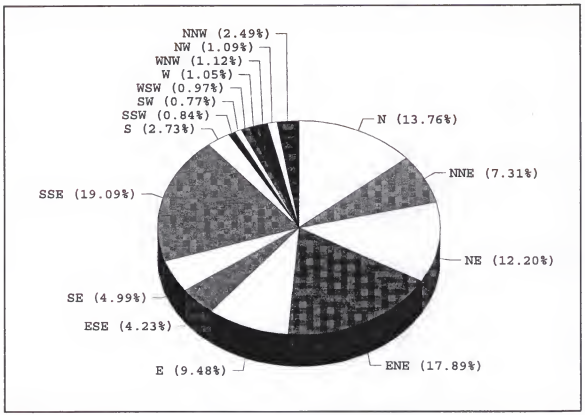


Figure 2-3 Pinellas Plant Cumulative Population Distribution

Lucie Unit 1 are latitude 27° 20' 58" north and longitude 80° 14' 48" west. Approximately 300 feet to the south of St. Lucie Unit 1 is FPL's St. Lucie Unit 2. St. Lucie Unit 2 is located at latitude 27° 20' 55" north and longitude 80° 14' 47" west. The Universal Transverse Mercator (UTM) coordinates for the midpoint (supplied by FPL) are 3,025,173 meters north and 574,326 meters east.

The eastern boundary of the site is the Atlantic Ocean and the western boundary is the Indian River, a tidal lagoon. Other prominent natural features within 50 miles of the site include Lake Okeechobee, 30 miles to the west-southwest of the site and a portion of the Everglades approximately 24 miles to the south of the site. Figure 2-4 shows the

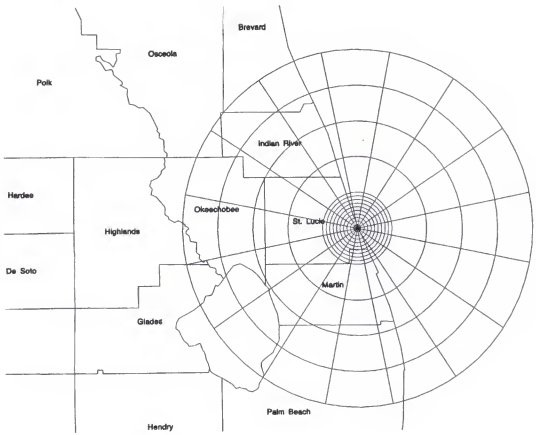


Figure 2-4 Sector-Segment Grid for St. Lucie Positioned Over County Boundaries

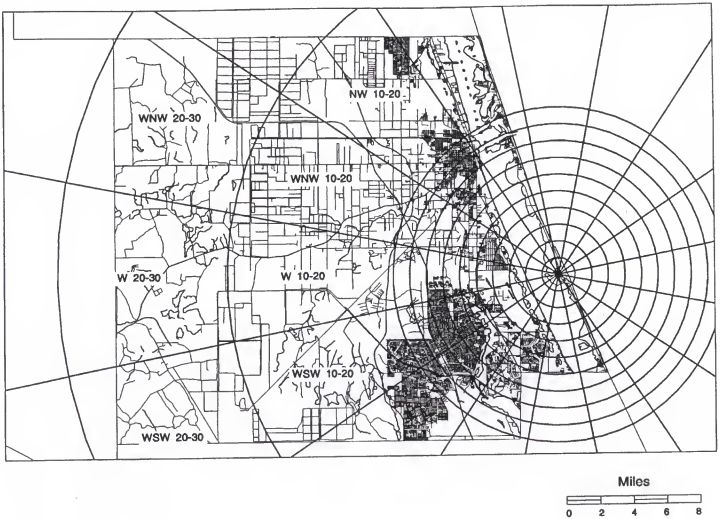


Figure 2-5 St. Lucie County with Sector-Segments

site in relation to the region within 50 miles. Figure 2-5 also shows the site with the 50-mile grid positioned over county boundaries. Figure 2-6 displays the level of census detail with a close-up view of the first 10 radial miles of the grid over census block boundaries.

Prominent cities within 10 miles of the site include Fort Pierce, approximately 7 miles to the northwest of the site on the mainland; Port St. Lucie, 4.5 miles to the west-southwest; and Stuart, 8 miles to the south. The largest urbanized area within 50 miles of the site is West Palm Beach, located 36 miles to the south-southeast.

Transportation corridors within 5 miles of the site include U.S. Highway 1 (US 1); State Roads (SR) A1A, 712 and 707: the Florida East Coast Railroad; the Atlantic Ocean and the Intracoastal Waterway which is located in the Indian River. SR A1A, the major north-south route on Hutchinson Island, traverses FPL's property to the east of St. Lucie Units 1 and 2.

Site Description

FPL owns approximately 1,132 acres of land on Hutchinson Island. The site is generally flat and has dense vegetation characteristic of Florida coastal mangrove swamps. At the ocean shore, the land rises slightly to a dune or ridge approximately 19 feet above mean sea level. The area pre-empted by the plant is about 300 acres, or approximately 27 percent of the total land owned by FPL. There are no industrial, commercial, institutional, or residential structures within the plant area. SR A1A traverses FPL's property approximately 1,000 feet east of the St. Lucie Plant. The radius of the exclusion area is 0.97 miles from the center of the St. Lucie Plant. The low population zone includes that area within 1 mile of the center of the St. Lucie Plant.

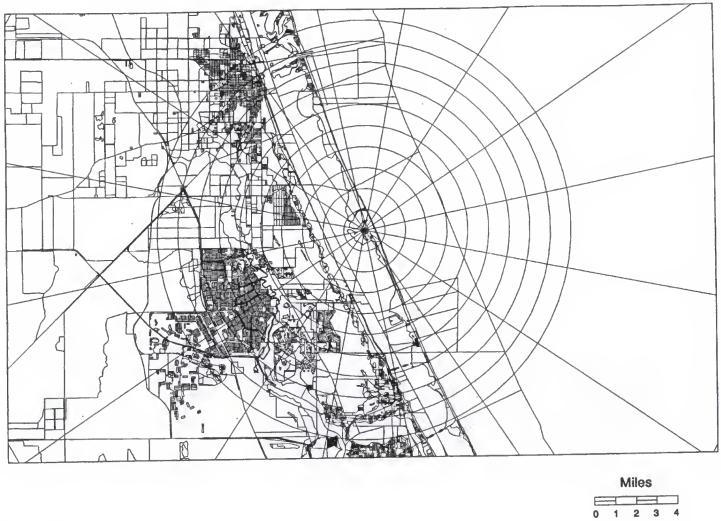


Figure 2-6 Sector-Segment Grid Over Census Blocks Within the 10-Mile Radius

Population Within 10 Miles

There are an estimated 132,958 people (1990 data) who reside within 10 miles of the St. Lucie Nuclear Power Plant. The population is concentrated in the cities of Fort Pierce and Port St. Lucie, both in St. Lucie County. Most of the area within 10 miles of the plant is in St. Lucie County. A small portion of the 10-mile area south of the plant is in Martin County. Hutchinson Island extends along the entire eastern boundary of the St. Lucie Plant 10-mile area, as the remaining area east of the island is water. There are 2 portions of Hutchinson Island, the north island and the south island. A small portion of the north island is within 10 miles of the plant, in sector NNW 5-10. The majority of the south island is within the plant's 10-mile radius in both St. Lucie and Martin Counties. A number of large developments have been built on the island, most of these being south of the plant.

Cities, Towns and Settlements

The largest population centers within 10 miles are the cities of Fort Pierce and Port St. Lucie. The city of Fort Pierce lies northwest and north-northwest of the St. Lucie Plant. Most of its area is located between 5 and 10 miles of the plant to the northwest, except for the community of Collins Park Estates which extends to approximately 3 miles from the plant in the west-northwest and western sectors. Sector NNW includes primarily that area of Fort Pierce which is on Hutchinson Island. This portion of Hutchinson Island has been extensively developed and includes a mixture of single-family homes, mobile homes, multi-family condominium developments and a number of tourist accommodations. The 1990 resident population in Fort Pierce was 36,830. The city has experienced a population growth of 9% over its 1980

population of 33,802. This represents an average annual increase of about 0.9%.

The city of Port St. Lucie lies west, southwest and south of the plant. Its city boundaries range from about 3 miles to 13 miles from the plant. Port St. Lucie encompasses a relatively large portion of the area within 10 miles of the plant. In 1980, this city's population was only 14,690. Since that time, the number of persons residing in Port St. Lucie increased to 55,866 in 1990, an increase of 280.3%. Much of this growth has occurred within 10 miles of the St. Lucie Plant. Development of the area between 5 and 10 miles has been primarily in the form of single-family homes. Most multi-family developments have remained closer to the plant, off U.S. Highway 1.

The following is a breakdown of development according to 1990 census data in terms of housing units, housing unit density per square mile and significant communities or residential features by sector-segment. Tables 2-8, 2-9 and 2-10 give housing unit densities per square mile of land. Housing unit density per area of land is used to eliminate dilution effects of large percentages of water in some sectors. Areas of less than 1 square mile are not considered, because densities could be artificially inflated by using a fractional denominator in the density quotient. First, consider the southern sectors SE through SSW.

The SE sector approaching the 2-mile radius contains the first of the condominium developments south of the plant on Hutchinson Island. The rest of the island south of the plant is situated within the SSE sector. The most intense development lies from 4 to 6 miles from the plant where 4,563 housing units have been constructed, however numerous multi-story condominiums and other enterprises span the length of the

Table 2-8 St. Lucie Housing Unit Densities to 5 Miles

| 1990 HOUSING UNIT DENSITY BY SECTOR-SEGMENT (PER SQUARE MILE OF LAND) | | | | | |
|---|---|---|---|---|---|
| | DISTANCE (MILES) | | | | |
| DIRECTION | 0-1 | 1-2 | 2-3 | 3-4 | 4-5 |
| N | 0.00 | **** | **** | 0.00 | 0.00 |
| NNE | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| NE | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| ENE | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| E | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| ESE | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| SE | 0.00 | **** | **** | 0.00 | 0.00 |
| SSE | 0.00 | **** | **** | **** | **** |
| S | 0.00 | 0.00 | 0.00 | **** | **** |
| SSW | 0.00 | 0.00 | **** | 251.88 | 514.57 |
| SW | 0.00 | **** | **** | 174.08 | 728.44 |
| WSW | 0.00 | **** | **** | 132.25 | 1,337.24 |
| W | 0.00 | **** | **** | 361.04 | 234.17 |
| WNW | 0.00 | 0.00 | **** | 444.57 | 688.64 |
| NW | 0.00 | **** | 0.00 | **** | **** |
| NNW | 0.00 | **** | **** | **** | **** |
| Average by Annulus | 0.00 | 0.00 | 0.00 | 272.76 | 700.61 |

**** Land area of less than 1 square mile would yield inflated densities.

southern island. These include Sand Dollar Villas, Dune Walk, Admiral, Island Dunes I-II, Ocean Towers I-II, Islandia I-II, Empress, Princess, Island Beach Resort, Oceana II North, Oceana I North, Miramar "Royal", Miramar I-II, Sea Winds, Atlantis "B", Atlantis "A", Atlantis III, Ocean Rise, Outdoor Resorts (Nettles Island), Island Club, Oceana South, Ocean South II, Turtle Reef Club, Turtle Reef, Beach Club Colony, Ocean Dunes, Holiday Out, Windmill Village and Villa Del Sol. Development continues in a similar pattern in Martin County with The Dunes, Jensen Beach Club, Indian River Point, Fairwinds Cove, Green Turtle Cove, Seaside, Joe's Point, Sandpebble (Ocean), Sandpebble (River), Ocean View, Ocean View (River), Beachwood Villas, Buttonwood, Shore Villas, Maritimes East, Maritimes West, Hutchinson House East, Hutchinson House West, Angler's Cove and Little Ocean Place. Many of the aforementioned condominium

Table 2-9 St. Lucie Housing Unit Densities to 10 Miles

| 1990 HOUSING UNIT DENSITY BY SECTOR-SEGMENT (PER SQUARE MILE OF LAND) | | | | | |
|---|---|---|---|---|---|
| | DISTANCE (MILES) | | | | |
| DIRECTION | 5-6 | 6-7 | 7-8 | 8-9 | 9-10 |
| N | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| NNE | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| NE | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| ENE | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| E | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| ESE | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| SE | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| SSE | **** | **** | **** | **** | **** |
| S | 293.44 | 378.39 | 780.24 | 1,158.02 | **** |
| SSW | 590.46 | 396.30 | 352.08 | 557.20 | 164.30 |
| SW | 727.41 | 813.95 | 700.98 | 310.83 | 363.63 |
| WSW | 1,036.28 | 1,167.46 | 815.26 | 246.65 | 206.08 |
| W | 221.97 | 166.90 | 159.21 | 7.61 | 0.55 |
| WNW | 525.80 | 201.31 | 104.86 | 98.23 | 65.03 |
| NW | 659.60 | 911.93 | 1,151.46 | 1,444.39 | 1,254.41 |
| NNW | **** | 822.55 | **** | **** | 300.82 |
| Average by Annulus | 579.28 | 607.35 | 580.58 | 546.13 | 336.40 |

**** Land area of less than 1 square mile would yield inflated densities.

communities are used on a seasonal basis. The island continues beyond the 10-mile radius. This sector also contains the entirety of the peninsular community of Sewall's Point, but only the northern, less-developed portion lies within the 10-mile radius.

Development of interest in the S sector begins in the 4-5 segment west of the Green River Parkway continuing into segment 5-6 north of the county line in St. Lucie County and into Martin County along the St. Lucie River shore involving the northern part of Pineapple Bluff. That development is the most prominent in segment 6-7 around Savanna Road. Proceeding into segment 7-8 Savanna Road and Jensen Beach Boulevard (SR 732) service Sugar Hill, Holly Creek, Timberwick, Pinechest Lakes and

Table 2-10 St. Lucie Housing Unit Densities to 50 Miles

| 1990 HOUSING UNIT DENSITY BY SECTOR-SEGMENT (PER SQUARE MILE OF LAND) | | | | |
|---|---|---|---|---|
| | DISTANCE (MILES) | | | |
| DIRECTION | 10-20 | 20-30 | 30-40 | 40-50 |
| N | 0.00 | 0.00 | 0.00 | 0.00 |
| NNE | 0.00 | 0.00 | 0.00 | 0.00 |
| NE | 0.00 | 0.00 | 0.00 | 0.00 |
| ENE | 0.00 | 0.00 | 0.00 | 0.00 |
| E | 0.00 | 0.00 | 0.00 | 0.00 |
| ESE | 0.00 | 0.00 | 0.00 | 0.00 |
| SE | 0.00 | 0.00 | 0.00 | 0.00 |
| SSE | 548.48 | 394.14 | 888.79 | 1,653.38 |
| S | 361.21 | 20.21 | 49.67 | 422.43 |
| SSW | 58.54 | 19.17 | 0.55 | 2.82 |
| SW | 22.02 | 3.87 | 2.48 | 221.62 |
| WSW | 17.88 | 2.74 | 145.13 | 15.08 |
| W | 5.68 | 1.36 | 27.31 | 7.25 |
| WNW | 12.31 | 0.51 | 2.45 | 1.41 |
| NW | 117.63 | 53.21 | 8.09 | 0.83 |
| NNW | 516.40 | 368.45 | 138.04 | 118.28 |
| Average by Annulus | 184.46 | 95.96 | 140.28 | 271.45 |

**** Land area of less than 1 square mile would yield inflated densities.

Ocean Breeze Park (mobile home community) in Jensen Beach. SR 732 also provides the focal point for the most numerous (3,873 units) and dense (1,158 units/mile$^2$) housing unit construction in this sector within the 10-mile limit. This occurs in the 8-9 segment which includes Vista Del Sol, Pine Lake Village MHP, Jensen Park Estates, Savannas Club, Tropical Acres MHP and Beacon 21. The last segment (9-10) is comprised of 800 Place Garden Villas, Sunset Cove, Harpers Landing, as well as development around SR 707 in the Rio area.

The SW sector incorporates part of the Harris Subdivision and Riverview Heights plus residential areas south of Walton Road and west of the Green River Parkway in the 3-4 segment. Growth in segment 4-5 has arisen along Melaleuca Boulevard. Community growth west of the

Florida Turnpike and east of Lennard Road supplies the greatest number and highest density of housing units in the 10-mile radius of this sector from 5 to 6 miles from the plant with 1,295 units and 590 units per square mile.

Development continues in segment 7-8 along US 1 and Port St. Lucie Boulevard in St. Lucie County, however sparse growth in the Martin County constituent fuels a decrease in both total housing units and their density. This dual decreasing trend continues to the 10-mile limit. Housing is concentrated along Westmoreland Boulevard in segment 7-8 and is sporadic in Martin County. Segment 8-9 features development along the northern shore of the North Fork of the St. Lucie River in both St. Lucie and Martin Counties. The residential area of the terminal segment lies south and west of Harbour Ridge Country Club adjacent to Southbend Boulevard.

There is limited development in the SW sector until segment 3-4 where Savanna Club, an adult manufactured community, is located. In general, development in this sector is concentrated around major roadways. Specifically, US 1 and Walton Road in segment 4-5; US 1, Lyngate Drive and Midport Road in segment 5-6; Midport Road, Port St. Lucie Boulevard and north of the North Fork of the St. Lucie River in segment 6-7; Whitmore Drive and Port St. Lucie Boulevard north of the North Fork of the St. Lucie River and east of Westmoreland Boulevard in the southern portion of segment 7-8; Port St. Lucie Boulevard, the Florida Turnpike and the development of Vikings Lookout in segment 8-9; and Port St. Lucie Boulevard, the Florida Turnpike, Savage Boulevard, Tulip Boulevard and Darwin Boulevard in segment 9-10. The number of housing units and their density (per square mile of land with water subtracted) peak in segment 6-7 at 2,023 units and 814 units per square mile of land.

Sector WSW contains parts of Spanish Lakes MHP and St. Lucie Gardens in segment 3-4 and quickly reaches a maximum density of 1,337 units per square mile of land in the next 1-mile segment with the inclusion of more of Spanish Lakes MHP and St. Lucie Gardens. Hidden River Estates, River Park and River Park Unit 2 comprise segment 5-6. There is widespread residential concentration around Floresta Drive and Prima Vista Boulevard in segment 6-7 where the sector high of 2,951 housing units are situated. Prima Vista Boulevard and Whitmore Drive are the hubs of development in segment 7-8. The last 2 segments contain parts of St. Lucie West. Most of the development in segment 8-9 is east of the Florida Turnpike and south of Juliet Avenue. Approximately equal development exists in the last segment, primarily south of Juliet Avenue.

Development in sector W is not noteworthy until segment 3-4 where both the number of housing units and their density abruptly reach maxima for the sector within 10 miles of the plant. These are 483 units and 361 units per square mile, respectively. Together with segment 4-5, These segments make up the heart of the White City area and include parts of Indian River Estates, Lexington Square, River's Edge and The Woodlands. Segment 5-6 includes part of Driftwood Manor in the north and Rivers Edge in the south. Continuing to the next segment (6-7), some of Lucy Acres in the north and extensive development west of St. James Drive as well as north and south of Bayshore Boulevard make the main contributions, yet a decreasing trend in units and densities continues throughout this and the remainder of the sector. The residential pattern established in the previous segment prevails in segment 7-8 predominantly east of the Florida Turnpike.

Development drops off precipitously in the remainder of the radius of interest, with much of it located within the loop of the Torino Parkway.

In the NNW sector, significant development is encountered in the 3-4 segment where portions of Indian River Estates in White City are located. Gator Trace and St. James Park, both of Ft. Pierce, join Indian River Estates to comprise the residential tracts of segment 4-5. This is the most densely developed segment of the sector containing 689 housing units per square mile and 1,172 total units. More of St. James Park along with parts of Driftwood Manor and Timber Ridge Estates make up segment 5-6. Midway Road and Oleander Boulevard provide development loci. Segment 6-7 includes pieces of Canoe Creek, River Hammock and Lucy Acres with development focused on Sunrise Boulevard. Sweetwater and Raintree Forest are the main communities in segment 7-8. In addition, there is some development around Midway Road. About the same level of development continues into segment 8-9 with Lost Tree Estates, River Oaks Estates, Thousand Pines Estates and more all around Edwards Road. The last mile increment of the 10-mile radius is less developed, but features Westglen and Briargate.

The NW sector contains almost all of incorporated Ft. Pierce. As before, segment 4-5 includes part of St. James Park. Development in segment 5-6 is primarily east of US 1. Most of the development in segment 6-7 is in Oleander Gardens east of US 1. In segment 7-8, development is predominantly north of Edwards Road. It is extensive west of US 1 to Oleander Boulevard (SR 605) and continues west of SR 605 to Sunrise Boulevard and beyond except for the area occupied by the Lawnwood Recreation Complex. The next segment, 8-9, exhibits the peaks of 4,749 housing units and 1,444 units per square mile for this sector. This is due to concentrated development in Ft. Pierce around Okeechobee

Road (SR 70); the east-west thoroughfares of Orange Avenue, Delaware Avenue and Virginia Avenue; and 25th Street (SR 615) running north-south. Tucker Terrace and Lawnwood are the largest communities. The final segment decreases in housing unit density with the majority of units built along Orange Avenue and 25th Street.

The final sector under consideration is NNW. This sector encompasses the majority of Hutchinson Island north of the plant. Segment 6-7 has development east of A1A on the island including Surfside Harbor. Development in segment 7-8 stretches across the island from Jennings Cove on the lagoon side to Surfside and Tropical Beach on the Atlantic side. Segment 8-9 contains the maximum of 1,637 housing units which can be ascribed to Thumb Point, Bayshore Estates and Causeway Mobile Home Park on the island with little development on the small section of mainland in the segment. All of the South Bridge is within the segment. Most of the North Bridge is in segment 9-10. Residential development lies on the southern tip of the north island between Coral Cove and Ft. Pierce Inlet Recreation Area. There is little development on the mainland inclusion.

Population by Annular Sectors

The most heavily populated annular sectors are those which cover the towns and developments mentioned above. The most heavily populated sector in the 10-mile radius is annular sector NW 5-10 with a total of 35,265 residents, which includes much of the City of Fort Pierce, with a 1990 population of 36,830 residents.

Population by Annuli

The annulus between 5 and 10 miles of the St. Lucie Plant is more densely populated than the area within 5 miles. A total of 116,230 live between 5 and 10 miles of the plant (1990 data). Inside 5 miles, there

are about 16,728 residents. Within 2 miles of the St. Lucie Plant, there is an estimated total of 154 residents. The entire area within 1 mile of the plant is owned by FPL and is included in the exclusion area and low population zone. Much of the area in the one- to two-mile annulus is water.

Population by Sectors

The most populous sector within 10 miles of the St. Lucie Plant is the NW sector which, because of the large concentration of resident population in the city of Fort Pierce, contains 35,636 persons (1990). The second most heavily populated sector is WSW, which has 23,778 persons and includes part of Port St. Lucie. The adjacent sector, SW, is the third highest with 21,019 residents in 1990.

Population Between 10 and 50 Miles

Tables 2-11, 2-12 and 2-13 show the distribution of the calculated 1990 population between 10 and 50 miles of the St. Lucie Plant; and Figure 2-5 presents that information figuratively. The calculated 1990 population is 659,411 persons and represents 83.2 percent of the total population within 50 miles of the plant. This population is confined to sectors SSE through NNW since sectors N through SE, beyond the 10 mile radius, include only the Atlantic Ocean. The major concentration of population occurs in annular sector SSE 40-50, which includes portions of the city of West Palm Beach, Palm Springs, Haverhill, Greenacres City, Royal Palm Beach and Wellington. West Palm Beach is the northern limit of the Florida Gold Coast development extending north from Miami through Dade and Broward Counties into Palm Beach County. The 151,310 residents in annular sector S 40-50 live on approximately 174.5 square

Table 2-11 St. Lucie Resident Populations to 5 Miles

| 1990 RESIDENT POPULATIONS BY SECTOR-SEGMENT | | | | | |
|---|---|---|---|---|---|
| | DISTANCE (MILES) | | | | |
| DIRECTION | 0-1 | 1-2 | 2-3 | 3-4 | 4-5 |
| N | 0 | 0 | 0 | 0 | 0 |
| NNE | 0 | 0 | 0 | 0 | 0 |
| NE | 0 | 0 | 0 | 0 | 0 |
| ENE | 0 | 0 | 0 | 0 | 0 |
| E | 0 | 0 | 0 | 0 | 0 |
| ESE | 0 | 0 | 0 | 0 | 0 |
| SE | 0 | 118 | 0 | 0 | 0 |
| SSE | 0 | 0 | 0 | 197 | 976 |
| S | 0 | 0 | 0 | 196 | 443 |
| SSW | 0 | 0 | 38 | 710 | 2,100 |
| SW | 0 | 10 | 124 | 339 | 1,998 |
| WSW | 0 | 20 | 134 | 245 | 2,808 |
| W | 0 | 6 | 90 | 1,292 | 756 |
| WNW | 0 | 0 | 45 | 1,218 | 2,490 |
| NW | 0 | 0 | 0 | 23 | 348 |
| NNW | 0 | 0 | 0 | 0 | 0 |
| Total by Annulus | 0 | 154 | 431 | 4,220 | 11,919 |

miles of land (there are only 2.1 square miles of land in this sector-segment). Annular sectors SSE 40-50 and SSE 30-40 have the second and third highest populations , respectively and reflect that Palm Beach County is more highly developed than any other part of the region. Of the total 863,518 residents of Palm Beach County in 1990, about 44 percent lived within 50 miles of the St. Lucie Plant.

Cities and Towns Within 50 Miles

Table 2-14 lists towns, cities and communities within 50 miles of St. Lucie Plant with a 1990 population of more than 5,000, persons. There are 12 towns with a population of more than 10,000, the largest of which is West Palm Beach, with a 1990 population of 67,643. The second largest is the city of Port St. Lucie, with 55,866 persons in 1990; the

Table 2-12 St. Lucie Resident Populations to 10 Miles

| 1990 RESIDENT POPULATIONS BY SECTOR-SEGMENT | | | | | |
|---|---|---|---|---|---|
| | DISTANCE (MILES) | | | | |
| DIRECTION | 5-6 | 6-7 | 7-8 | 8-9 | 9-10 |
| N | 0 | 0 | 0 | 0 | 0 |
| NNE | 0 | 0 | 0 | 0 | 0 |
| NE | 0 | 0 | 0 | 0 | 0 |
| ENE | 0 | 0 | 0 | 0 | 0 |
| E | 0 | 0 | 0 | 0 | 0 |
| ESE | 0 | 0 | 0 | 0 | 0 |
| SE | 0 | 0 | 0 | 0 | 0 |
| SSE | 2,288 | 516 | 421 | 769 | 443 |
| S | 936 | 1,531 | 3,931 | 6,874 | 1,855 |
| SSW | 2,662 | 2,223 | 1,790 | 1,525 | 695 |
| SW | 3,154 | 5,104 | 4,554 | 2,402 | 3,334 |
| WSW | 3,966 | 6,819 | 6,054 | 1,796 | 1,936 |
| W | 1,247 | 1,120 | 1,161 | 37 | 5 |
| WNW | 2,310 | 1,378 | 801 | 819 | 612 |
| NW | 1,055 | 3,560 | 6,528 | 10,920 | 13,202 |
| NNW | 0 | 624 | 724 | 1,988 | 565 |
| Total by Annulus | 17,618 | 22,875 | 25,964 | 27,130 | 22,647 |

third largest is Fort Pierce with 36,830 persons; and the fourth largest is Riviera Beach in Palm Beach County with 27,639 persons. Of the 12 largest towns, 7 are in Palm Beach County. In addition to West Palm Beach and Riviera Beach, the 7 include North Palm Beach (11,343 persons), Jupiter (24,986 persons), Greenacres City (18,683 persons), Royal Palm Beach (14,589 persons) and Palm Beach Gardens (22,965 persons). Stuart, the largest city in Martin County, has a 1990 population of 11,936.

Of the 4 towns with populations between 5,000 and 10,000, all are within Palm Beach County. These include the towns of Palm Beach, with 9,814 persons; Lake Park, with 9,763 persons; and Palm Springs, with 9,763 persons. Pahokee, with a 1990 population of 6,822, is also in

Palm Beach County, but is located in the northwestern quarter of the county, on the shore of Lake Okeechobee.

Table 2-13 St. Lucie Resident Populations to 50 Miles

| 1990 RESIDENT POPULATIONS BY SECTOR-SEGMENT | | | | |
|---|---|---|---|---|
| | DISTANCE (MILES) | | | |
| DIRECTION | 10-20 | 20-30 | 30-40 | 40-50 |
| N | 0 | 0 | 0 | 0 |
| NNE | 0 | 0 | 0 | 0 |
| NE | 0 | 0 | 0 | 0 |
| ENE | 0 | 0 | 0 | 0 |
| E | 0 | 0 | 0 | 0 |
| ESE | 0 | 0 | 0 | 0 |
| SE | 0 | 0 | 0 | 0 |
| SSE | 20,115 | 32,205 | 78,634 | 114,025 |
| S | 38,024 | 5,778 | 13,255 | 151,310 |
| SSW | 7,361 | 5,241 | 178 | 1,398 |
| SW | 4,434 | 1,141 | 246 | 8,810 |
| WSW | 3,012 | 740 | 15,196 | 1,824 |
| W | 698 | 381 | 10,176 | 2,478 |
| WNW | 2,078 | 143 | 811 | 494 |
| NW | 17,967 | 8,450 | 2,882 | 208 |
| NNW | 25,131 | 44,776 | 19,189 | 20,622 |
| Total by Annulus | 118,820 | 98,855 | 140,567 | 301,169 |

Population by Annular Sectors

The most heavily populated annular sectors between 10 and 50 miles from the St. Lucie Plant are those which encompass the cities and towns with the greatest populations. The most populous annular sector, S 40-50, includes Greenacres City (18,683 persons in 1990), Haverhill (1,058 persons in 1990), Palm Springs (9,763 persons in 1990) and Royal Palm Beach (14,589 persons in 1990).

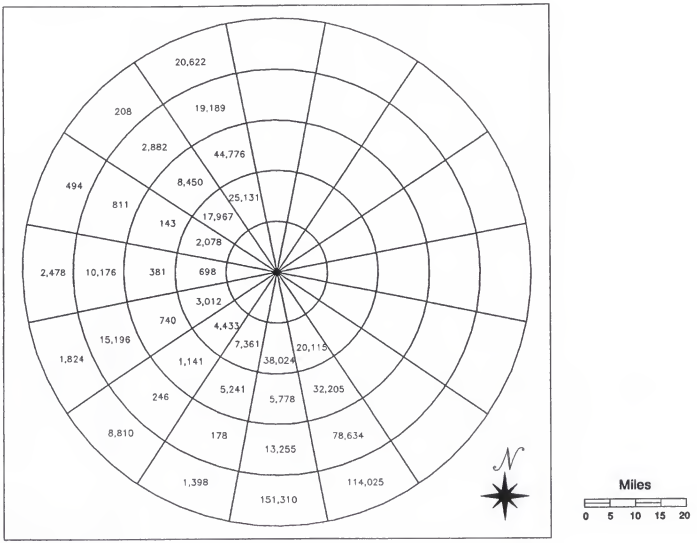


Figure 2-5 Resident Population from 10 to 50 Miles for the St. Lucie Plant

The second most populous annular sector is SSE 40-50, which includes portions of Palm Beach (9,814 persons in 1990), Riviera Beach (27,639), West Palm Beach (67,643), as well as Cloud Lake (121), Glen Ridge (207), Lake Clarke Shores (3,364) and Mangonia Park (1,453).

Table 2-14 Towns, Cities and Communities Within 50 Miles of the Plant

| Communities of over 10,000 Persons | | | | |
|---|---|---|---|---|
| City or Town | County | 1970 Population | 1980 Population | 1990 Population |
| West Palm Beach | Palm Beach | 57,375 | 63,305 | 67,643 |
| Port St. Lucie | St. Lucie | 330 | 14,690 | 55,866 |
| Fort Pierce | St. Lucie | 29,721 | 33,802 | 36,830 |
| Riviera Beach | Palm Beach | 21,401 | 26,489 | 27,639 |
| Jupiter | Palm Beach | 3,316 | 9,868 | 24,986 |
| Palm Beach Gardens | Palm Beach | 6,102 | 14,407 | 22,965 |
| Greenacres City | Palm Beach | 1,731 | 8,780 | 18,683 |
| Vero Beach | Indian River | 11,908 | 16,176 | 17,350 |
| Royal Palm Beach | Palm Beach | | 3,423 | 14,589 |
| Stuart | Martin | 9,086 | 9,467 | 11,936 |
| North Palm Beach | Palm Beach | 9,035 | 11,344 | 11,343 |
| Sebastian | Indian River | | 2,831 | 10,205 |
| Communities of between 5,000 and 10,000 Persons | | | | |
| City or Town | County | 1970 Population | 1980 Population | 1990 Population |
| Palm Beach | Palm Beach | 9,086 | 9,729 | 9,814 |
| Palm Springs | Palm Beach | 4,340 | 8,166 | 9,763 |
| Pahokee | Palm Beach | 5,663 | 6,346 | 6,822 |
| Lake Park | Palm Beach | 6,993 | 6,909 | 6,704 |

The information in this table is based upon "Florida Population: Census Summary 1990" April 1991. Bureau of Economic and Business Research, University of Florida.

The third most populous annular sector between 10 and 50 miles of the St. Lucie Plant lies north of West Palm Beach on the Atlantic Coast (SSE 30-40). Although its land area is only 49.33 square miles of the 137.31 total square miles allotted to the sector-segment, i.e. 35.9 percent, it includes Lake Park (6,704 persons in 1990), North Palm Beach (11,343), Juno Beach (2,121), as well as portions of Riviera Beach (27,639), Palm Beach Gardens (22,965), Palm Beach Shores (1,040) and the

town of Jupiter (24,986). When the preceding three annular sectors are combined, they comprise 52.2 percent of the total population between 10 and 50 miles of the St. Lucie Plant.

Population by Annuli

Populations of annuli between 10 and 50 miles of the St. Lucie Plant range in number of residents from the largest, with a total of 301,169 persons (the 40-50 mile annulus), to the smallest, with 98,855 persons (the 20-30 mile annulus). The annulus between 30 and 40 miles has the second largest population of 140,567, while the annulus between 10 and 20 miles contains 118,820 persons.

The 40-50 mile annulus has not only the largest population and the greatest overall area (approximately 1,588 square miles of total area, excluding the seven sectors over the Atlantic Ocean; or 1,091 square miles of land area), but also the highest population density in the region. The population density of the 40-50 mile annulus is 620 persons per square mile of land (an average of the densities of the 9 sector-segments). Eighty-eight percent of the population is located on 22 percent of the total annulus area, in sectors SSE and S, which include West Palm Beach and environs.

Population by Sectors

The most populous sectors between 10 and 50 miles of the St. Lucie Plant are those which cover the West Palm Beach area and the Atlantic Coast. Sectors SSE and S have estimated 1990 populations of 244,979 and 208,367, respectively; and densities ranging from 792 to 3,188 persons per square mile of land and from 61 to 867 persons per square mile of land, respectively. Sector NNW has a population of 109,718 and densities from 238 to 869 persons per square mile of land; sector NW, the next one inland, has a total population of 29,507 and a density

range of 1.2 to 311 persons per square mile of land. The 5 remaining sectors have densities which range from 1.3 to 623 persons per square mile of land. The sparseness of population in the 5 interior sectors can be attributed to extensive acreage covered by wetlands and surface water (Lake Okeechobee), inaccessibility to population centers and the extent of range and cropland.

## Summary

This type of detailed information by annular rings, sector and sector-segment is made possible by use of the GIS as previously described. The sector-segment grid is constructed using the analysis tools of the GIS so that each sector-segments a separate analytical entity. This allows the user to query each entity for summary statistics of the geographic attributes, e.g. land area, water area, housing units and population. The grid is placed over the TIGER street map features (highways, waterways, city roads, etc.) To provide information about the specific communities. This is done by positioning the cursor on the object of interest (line or point) to access its name. Further enhancements include the use of more elaborate base maps to derive that information and reference to paper maps using the GIS map as a spatial reference.

CHAPTER 3
DEVELOPMENT OF POPULATION ESTIMATES AND PROJECTIONS

## Estimate and Projection Concepts

The concepts of population estimation and population projection are similar in that they each involve generating a number intended to indicate the size of the population of a given area at a specific point in time. Both population estimates and population projections make use of the basic demographic equation:

$$P_1 = P_0 + B - D + I - O$$

where

| | |
|---|---|
| $P_1$ | is the current population |
| $P_0$ | is the population at a previous point in time |
| B | is the number of births since $P_0$ |
| D | is the number of deaths since $P_0$ |
| I | is the number of inmigrants since $P_0$ |
| O | is the number of outmigrants since $P_0$ |

This equation expresses the population at any given point in time as a function of the population at a previous point in time and considers the amount of natural change and net migration during the interim. Natural change (usually an increase) is the number of births minus the number of deaths in the study area, and net migration is the number of inmigrants to the area minus the number of outmigrants away from the area. The sources of vital registration and migration data were discussed previously.

A population estimate refers to the size of the population of an area in the recent past, while a population projection refers to the

size of the population of an area at some point in the future. Population estimates are generally based on direct components of population change, such as the actual number of births and deaths occurring between the date of the previous population and the date of the estimate; in the absence of direct data, a population estimate will be based on symptomatic indicators of the components of population change. Symptomatic data are those data that move in concert with the data of direct interest. For example, a change in the school enrollments or the number of motor vehicles being registered may serve as symptomatic indicators of migration. Since population projections refer to the size of the population at a point in the future, they cannot be based on actual data comprising the components of population change. Population projections must in one form or another be based on the extension of either current or expected population trends into the future.

The only "true" population counts are available in decennial census years, e.g. 1970, 1980, 1990, etc. Of course, the most recent was the 1990 census which serves as the base for present and future projections until the year 2000 census is taken and published. Population projection methodologies at the county level have been developed by various State agencies. In 1967, the Federal-State Cooperative Program for Local Population Estimates (FSCPE) was formed. Through member participation, the Bureau of the Census has an opportunity to evaluate the level of accuracy of the population estimates produced between censuses and to review the adequacy of specific procedures used to prepare the estimates.

## Methods of Estimation

The most widely accepted estimation methods are described in the subsections that follow. Methodological strategies such as proration techniques, simple ratio methods, the censal ratio method, and the vital rates method are relatively simple to calculate and have minimal data requirements. Alternative postcensal estimation strategies such as composite methods, the U.S. Bureau of the Census component method II, the ratio correlation method, the administrative records method, the cohort component method, and the housing unit method involve somewhat more complex calculations and have more extensive data requirements.

### Vital Rates Method

The census year ratio of symptomatic data to population for a given area is applied to a current value of that symptomatic data to estimate the area's population. The symptomatic data are births and deaths. First, births and deaths are converted to rates, the rates are then adjusted to the estimate date by changes in these rates for some broader area which can be measured over time. Next the rates are divided into the number of births and deaths on the estimate date to yield two population estimates. The two estimates are usually averaged (USBC, 1990c).

### Composite Method

Component, regression, and other methods are combined in an effort to maximize the potential of each. The total for all ages results from the sum of estimates for various age groups or age-sex groups.

### Administrative Records Method

Uses births, deaths, internal migration, and immigration from abroad. The internal migration component is derived from individual tax

return records supplied by the Internal Revenue Service. A migration rate is derived by comparing the addresses on two years of tax returns when the addresses are not the same. Data on Medicare enrollees estimates the population 65 years old and over. The estimates of net migration from the under 65 years of age tax base combined with independent estimates of the population 65 years old and older, inmates of institutions, college students in dormitories, military personnel living in barracks, and other components of population change are added to the previous year's estimate to yield an estimate of the total population.

Ratio-Correlation Method

Uses a multiple regression equation to relate changes in two or more data series to estimated changes in population (the dependent variable). All variables are expressed as the ratios of percentage shares of the subarea to the universe in the estimate year to the corresponding percentage shares for the base year (previous census year). A separate regression equation is prepared and is unique for each State. The underlying assumption of this method is that past relationships between the independent and dependent variables will persist into the current estimating period. Some of the independent variables currently used are school enrollment, housing units, Federal income tax returns and exemptions, births, and deaths. Other possibilities include utility data, labor force data, drivers licenses, etc.

The ratio correlation method is one of the more versatile postcensal population estimation techniques. The ratio correlation method produces postcensal population estimates through the use of a statistical procedure called linear regression. Linear regression in

its simplest case is a statistical technique that allows researchers to predict one variable based on observations of a second variable. The variable to be predicted is referred to as the dependent variable (Y), and the variable used to make the prediction is referred to as the independent variable (X). Previous observations of values of X and Y are used to determine the statistical relationship between the two variables; given a new observation of X, one can use the statistical model to predict the value of Y. Linear regression analysis results in the development of an equation for a straight line that best fits the set of previous observations. The basic equation for a linear regression model is as follows:

$$Y = b + mX$$

Where:
- Y is the predicted value of the dependent variable
- b is the Y intercept, the point where the prediction line will cross the Y axis
- m is the regression coefficient or slope of the regression line
- X is the known value of the independent variable

The first step in creating the linear regression model is to calculate the value of m, the regression coefficient. The value of m represents the ratio of the change in the value of Y, the dependent variable, for each unit change in X, the independent variable. A computational formula for b is as follows:

$$m = \frac{n \cdot (\sum XY) - (\sum X) \cdot (\sum Y)}{n \cdot (\sum X^2) - (\sum X)^2}$$

Where:
- n is the number of observations
- $(\sum XY)$ is the sum of each X value multiplied by the corresponding Y value
- $(\sum X)$ is the sum of the X values
- $(\sum Y)$ is the sum of the Y values
- $(\sum X^2)$ is the sum of each X value squared
- $(\sum X)^2$ is the square of the sum of the X values

In the simplest case, a single independent variable is used to predict the dependent variable; in more complex cases, a set of independent variables may be used to predict the dependent variable through a multiple regression model. The multiple regression equation takes the form displayed below.

$$Y = b + m1X1 + m2X2 + \ldots + mnXn$$

In the ratio correlation method of population estimation, a set of dependent variables thought to be symptomatic of population change is used to predict the population of the estimate area. The technique is referred to as the ratio correlation method because ratio transformations of the independent variables, rather than the raw data, are used in the analysis.

Component Method II

Uses birth and death statistics to measure net natural change, but (unlike the Administrative Records method) it relies on school enrollment figures to measure net migration. The population 65 and over is based on the change in Medicare enrollees.

Housing Unit Method

Uses building permits, demolitions, and electric utility connections to estimate changes in the number of housing units or households and converts this to population change. The housing unit method produces postcensal population estimates by equating population to the number of occupied housing units times the average number of persons per household. The housing unit method is one of the most frequently applied methods of postcensal population estimation and is the primary method of choice for subcounty population estimates. The

following formula represents the logic of the housing unit method for estimating current population:

$$P = (H_o \cdot O) + I$$

where P is the current population
Ho is the number of occupied housing units
O is the number of persons per household
I is the institutional population

Two major methodological strategies have evolved for calculating the number of occupied housing units in an area. One strategy involves an analysis of building permit and demolition data, while the other strategy involves an analysis of utility data usually based on electric utility meters. Both strategies begin with the number of housing units known to exist at a previous point in time. The most recent census provides an enumeration of occupied housing units and the average number of persons per household which serves as the base data for the housing unit method.

Cohort-Component Method

Requires separate assumptions for each of the components of population change: births, deaths, internal migration, and international migration. Each component is projected separately for each birth cohort by sex and race.

Comparison Results

The Vital Rates Method and the Composite Method were replaced with the Administrative Records Method after comparison of published projections with decennial census results (USBC, 1986). For most States, the Administrative Records Method has emerged as the single most effective procedure for making population estimates, although accuracy is strengthened when this method is averaged with independent estimates

based on the Ratio-Correlation Method. Other methods in widespread use by State agencies include Component Method II and some State-specific methods. Some States use unique methods averaged with one or more of the above methods. One State-specific method gaining increasing favor is the Housing Unit method. For a simple example, consider a publication (USBC, 1988) in which the Ratio-Correlation Method was used to determine the prediction equation for the State of Tennessee. The regression equation for the State as a whole is given by

$$Y = -0.015 + 0.397\,(X1) + 0.127\,(X2) + 0.085\,(X3) + 0.412\,(X4)$$

where X1 = Federal income tax returns,
X2 = Medicare enrollees,
X3 = resident deaths, and
X4 = school enrollment in grades 1 through 8.

Note that there are key differences between State level and county level projections.

Table 3-1 County population change

| REGION | % CHANGE (TOT. POP.) | % NET MIGRATION |
|---|---|---|
| Tennessee | 3.8 | 0.9 |
| Bradley Co. | 6.1 | 2.5 |
| Hamilton Co. | -1.2 | -3.9 |
| Meigs Co. | 4.8 | 2.1 |
| Polk Co. | 0.4 | -0.8 |

The projected total population of Hamilton County actually decreased as driven by outmigration, while other counties and the State as a whole experienced different rates of migration and growth. The discussions that follow continue to contrast these levels and demonstrate the superiority of county level predictions over larger as well as smaller geographic and/or population coverages.

The 1990 publication "Projections of the Population of States, by Age, Sex, and Race: 1989 to 2010" (USBC, 1990b), used a cohort-component method that required separate assumptions for each of the components of population change as previously described. In addition, the internal migration component was estimated in four different ways, each with its own alternative assumption. Although the cohort-component technique is a good way to make projections at the state level, it is not necessarily the best way to make projections at the county level. Many counties are so small that the number of persons in each age-sex category is inadequate for reliable cohort-component projections. Even more important, county growth patterns are so volatile that a single technique based on migration data from one time period may provide misleading results. More useful projections can be made if several different techniques and historical time periods are used.

It is often the objective to produce projections within a single State rather than a consistent set of projections involving several or all States. In such instances, examining population projections prepared by State agencies can be useful. These State-produced projections represent an alternative to the projections developed by the Census Bureau. The individual State projections can be based on an assortment of models that incorporate a wider range of variables and data, since each State is not required to produce a set of projections consistent with other States (USBC, 1990b). It is evident that when one examines any particular State agency's projections that the numerical results are meaningless without a clear statement of the methods and assumptions that produced the numbers. FSCPE participants use those methods that have statistically proved to generate the most accurate projections for their respective States.

The most important and complex component of population change for States, counties, and small geographical areas is internal migration. With fertility and mortality rates in the U.S. becoming stable at historically low levels, internal migration has assumed critical importance in determining the growth or decline of counties and States (USBC, 1990a). Trends in internal migration differ throughout the U.S. There is also an interdependency at the State level, i.e., since net internal migration is zero by definition at the national level, an increase in net internal migration in one State has to be offset by decreases elsewhere. Internal migration is sensitive to changes in economic conditions, therefore both rapid and sizeable changes in internal migration can be expected. This is why local bureaus that monitor economic conditions are valuable to the purpose of population projection over a small region, e.g., within the 50-mile radius of a nuclear power or weapons production facility.

Population size is also a major determinant of accuracy levels. Upon comparison of projections with decennial counts in both 1970 and 1980, smaller counties (less than 5,000 population) experienced larger errors (7 percent) while larger counties (greater than 100,000 population) had errors of about half that amount (USBC, 1986). Clearly, a sparsely populated county such as Polk County,TN is at risk for larger error. However, the average error for a 10-year projection is still not excessive. Since the counties surrounding Polk are more populous (Bradley, McMinn, and Monroe), the inclusion of their more accurately projected blocks in shared sector-segments will dilute the Polk error contribution. In addition, even Polk County contains numerous census blocks; an asset that enables the investigator to confidently place the "true" 1990 figures in the correct sector-segments.

In a number of counties specific adjustments must be made to the population before applying the techniques described above. This is done to account for special populations such as university students, military personnel, and prison inmates. Adjustments are made for counties in which these special populations account for a large proportion of the total population or where the special populations have moved counter to trends for the rest of the population. The researcher draws on local FSCPE member agencies with an intimate knowledge of these factors and the methods best-suited to dealing with them. The application of the predictive equations generated by these agencies is revealed in the next section.

Population projections for sector-segments are estimated by applying common growth rates (predictive equations) to all census blocks within a particular county; then the software assigns weighted fractions of projected block populations to construct sector-segment totals. Projections are tabulated at user-defined intervals, and can be displayed in a variety of graphical forms (bar, rotated bar, 3-D bar, etc.). The reliability of FSCPE member's projections coupled with the sector-segment assignment system should result in the most accurate delineations of projected population distributions available.

## Elementary Postcensal Methods of Population Estimation

### Proration Estimation

The proration method of population estimation usually is employed as a method to allocate the population of a larger geographic area among its smaller subareas. The method is applied by allocating a portion of the current estimated population of a larger geographic area to a smaller geographic area based on the ratio of the population of the

smaller area to the population of the larger area as measured in the latest census. The formula for the proration method is as follows:

$$P_{es} = \frac{P_{cs}}{P_{cl}} \cdot P_{el}$$

where

| | |
|---|---|
| P | is the population |
| e | is an estimated value |
| c | is a census result |
| s | is a geographic subarea |
| l | is the larger geographic area |

The proration method is extremely easy to apply, has minimal data requirements, and may be applied to any level of geography. Application of the proration method is based on the assumption that the proportional relationship of the smaller and larger areas has remained constant since the time of the last census.

Simple Ratio Procedure

The simple ratio procedure is one of several strategies of postcensal population estimation relying on symptomatic data indicators. Symptomatic data are those data that serve as symptoms of population change; ideally, these are data series that change in either a positive or inverse manner in direct proportion to population changes. Logical choices include variables that serve as direct components of population change, such as births and deaths, along with other variables that are more indirect correlates of population change, such as voter registrations, school enrollments, utility hookups, and housing starts.

The logic of producing a postcensal population estimate through the use of a simple ratio technique involves measuring the value of a data series at two points in time along with the size of the population at the earlier point in time. Most often the earlier point in time will be the most recent census year. Once the data series has been measured

at the second point in time, a ratio is calculated indicating the change in the data series using the earlier point in time as the base. The change observed in the data series serves as an indicator of change in population. Population change is then measured by multiplying the population observed at the earlier point in time by the value of the ratio for the data series. The result is a current population estimate. One of the strengths of the simple ratio procedure is that multiple indicators of population change (several data series) may be used. Postcensal population estimates are obtained by averaging the various ratios when multiple indicators are used in the simple ratio procedure.

The simple ratio procedure is a good choice when data for an area are limited. It may be applied to a variety of levels of geography, and it is a technique that bases a population estimate on changes in variables that one logically can assume are related to population change.

Censal Ratio Method

Postcensal population estimates produced through the censal ratio technique are based on changes in the ratio of the level of predictor variables to the base population. In most applications some form of adjustment is made based on national trends or some other source. When no adjustments to the ratios are made, the censal ratio technique is essentially the same as the simple ratio technique.

The chief advantage of the censal ratio technique over the simple ratio technique is the ability to adjust the ratio indicating the relationship between the predictor variable and the population size. Any data series will undergo periodic fluctuations due to a variety of factors, and the fluctuations are likely to be more extreme at the local level. The ability to adjust the local ratio based on trends evident in

a larger geographic area prevents the data fluctuations in the smaller area from being incorporated into the local population estimate.

Shyrock and Siegel (1973) suggest that it is wise to select variables with certain characteristics. Variables for which data are collected and reported at frequent intervals will be preferable to those that are measured less frequently. Variables that have a sufficiently large number of events occurring each year compared with the size of the population are preferable to those with a smaller number of occurrences. Predictor variables that are relatively stable over time or that change in a predictable fashion should be chosen. Variables subject to extreme fluctuation or variables that are especially sensitive to factors other than population size should be avoided.

## Population Projections

Population projections are similar to intercensal estimates which are discussed in detail in Chapter 4. Since the methodology is essentially the same, the reader is referred to that chapter for examples. Population projections typically are based on the assumed continuation of the current situation, or perhaps on a high growth scenario, or on a low-growth scenario. The general methodological strategies for producing population projections fall into four basic categories: (1) ratio allocation methods, (2) mathematical extrapolation methods, (3) econometric methods, and (4) cohort component methods. Ratio allocation and mathematical extrapolation methods are simpler strategies for producing population projections. Ratio allocation methods are used to allocate an existing population projection for a larger area among the subareas that comprise it. For example, a county population projection may be allocated among the various blocks that comprise it, resulting not only in the original projection for the

county but in individual projections for each block. Mathematical extrapolation methods involve the application of a selected growth rate into the future. Econometric methods and cohort component methods are somewhat more complex strategies for producing population projections. These methods not only require additional technical expertise but require a great deal more input data to complete the projection process. The simpler methods of projection have a wider range of application and may be used to produce population projections at almost any geographic level. Econometric and cohort component strategies have more extensive data demands and are less appropriate for smaller areas of geography. For highly detailed population projections by age, race, and sex, the clear choice is a cohort component method. Econometric methods result in some detail by age, but not as much as cohort component methods.

Results of any methodological strategy are likely to be better the shorter the length of the projection period. Population change is a dynamic process with change sometimes occurring in ways that are unanticipated. The longer the length of the projection period, the more time will be available for the unanticipated to occur. Population projections also tend to be more accurate for areas that have low levels of vital rates such as fertility and mortality.

Results of population projections will tend to be more accurate the larger the area of geography analyzed. Several factors operate to improve the results for large areas. Population processes are volatile at times, but subarea fluctuations tend to be averaged out as one moves to a larger area of geography. For example, unusually high rates of population growth in one area will often be matched by unusually low rates of growth in another. In addition as the area of geography increases, the problems due to migration decrease.

Ratio Allocation Methods

Ratio allocation methods of population projection provide a way to allocate or distribute the projected population of a larger area among its constituent subareas. Essentially, one is given a population projection for a geographic area made up of identifiable subareas and proceeds to allocate the population projected for a given date among the subareas based on a set of operating assumptions. One general strategy is to assume that the subarea's share of the total population at an earlier date will remain constant over the length of the projection period. The other general strategy is to assume that some pattern of changing share is more appropriate, and the subarea will be allocated an increasing or decreasing share of the total projected population. The ratio allocation technique is especially well suited to small area geographies that are not easily projected by other methods. One of the major disadvantages is that the projections resulting from a ratio allocation method are only as good as the population projection for the larger area. In addition, the subarea's share of the total population almost certainly will change over time, which among other things will mean that the ratio allocation technique will not be as practical for longer projection periods.

Mathematical Methods

The primary methods of mathematical population projection involve applications of mathematical extrapolation, which is the extension of a mathematically defined trend. There are two major types of methods of mathematical population projection, and each type has several variations. One major type involves mathematical extrapolation, or simply extending a mathematical trend into the future. The mathematical extrapolation methods are similar to the methods of intercensal

population estimation utilizing mathematical interpolation, and one may use the same three types of assumptions. That is, given a growth rate for a population, one may assume an arithmetic rate of growth or a geometric rate of growth, with geometric rates of growth expressed as a function of periodic compounding or continuous compounding. The second major type involves fitting a growth rate to a curve. Three frequently chosen alternatives are the Gompertz curve, the Pearl-Reed curve, and a modified logistic curve. In each case a mathematical constant is used to place an upper limit on the amount of growth that may take place during the projection period. Each of the three curves results in a growth pattern characteristic of a logistic curve which has the form of an elongated S-shaped pattern suggestive of a period of slow gradual growth followed by a period of rapid growth and then culminating in a final period of no growth.

Case Study: St. Lucie Estimates and Projections

The following study is an example of an in-depth population projection series accomplished using a GIS platform, and a new, modified projection technique developed in this effort. The new technique is a modified proration strategy as explained below. Instead of the typical state to county direct procedure, this technique creates new geographical zones (sector-segments), area-weights census blocks from different counties to the new zones, then applies proration from county level to block level through the DBMS. The census blocks that are prorated from different counties are then synthesized into new composite regions by the GIS. The result is a prorated allocated projection. This case study demonstrates the complexity required in projections for a nuclear power plant and the detailed data generated by the new

technique. The St. Lucie Plant projection is an expansion of the base population presented in Chapter 2. The technique is enhanced using multiple sources and types of information such as city and county planning offices and housing permits.

Methodology

It is recommended that the non-demographer obtain population data professionally prepared by a State agency such as Florida's Bureau of Economic and Business Research. These documents are inexpensive and updated periodically. Use the county-level populations for projection years to generate growth ratios. Create new fields in the population database for each projection year and apply the growth ratio for each county to its census blocks in the spreadsheet. Next, construct the sector-segment grid as before and query the selected regions for the new area-weighted projections. Each sector-segment projection will be a composite of those projected blocks from the various counties that comprise it. If a researcher has an abundance of demographic data (especially migration data) and does not wish to use the simple procedure designed for this study (above), then the techniques described previously should be addressed. However, once the projections are made for counties, the GIS procedure remains the same. The solution to the problem of saturation lies in whether or not one accepts the following assumptions. Assume that the highest population density for a census block in the base census year (e.g., 1990) is a maximum that the local population can not or does not wish to exceed. Next, assume that adjacent or nearby high-density blocks are the more desirable to attract excess populace. In the database write a script appropriate for the software of choice that calculates the projected block, subtracts the population that exceeds the density maximum, and forces that excess to

the second highest density block. Saturated blocks must be forced to stagnancy.

A second problem is that of null blocks. One can observe from the case study projections that regions with zero as a baseline can never grow using growth equations. In otherwise highly populated areas, such blocks usually involve water, farmland, industrial areas, etc., but this is not always the case. If the GIS user has good information about planned developments, the null regions can be "seeded" with families to generate future populations.

Projected Population

The population within 10 miles of the St. Lucie Plant is expected to increase by about 19% over the 5-year period between 1990 and 1995. The 1995 resident population is projected to be 157,625. The continued development of Port St. Lucie and Hutchinson Island are expected to be the largest contributors to this growth. The most significant population increase in Port St. Lucie will be attributed to St. Lucie West, a 4,600-acre tract of land to be developed by the J. White Corporation. St. Lucie West is situated in sectors WSW and W. Most of its area is between 5 and 10 miles of the plant, with some of it extending beyond the 10-mile radius. The St. Lucie West planned development also includes a spring training sports complex for the Mets which was completed in 1988, new college campuses for Indian River Community College and Barry University, public schools, an industrial park, office park, and a regional mall.

Development along unincorporated Hutchinson Island north of the plant has been slow and is projected to remain this way in the near future. However, this trend will probably change as the southern part of the island becomes more and more congested.

County planning officials have indicated that congestion of the bridges from the mainland to Hutchinson Island could restrict development. A bridge has been proposed which would cross the Indian River at SR 712 and link US 1, the Florida Turnpike, and Interstate 95 to Hutchinson Island. An additional river crossing would induce development on the island. However, it is uncertain when or where another river crossing will be constructed, because the waters of the Indian River in this area are part of an aquatic preserve.

Population Between 10 and 50 Miles

Projected population

Total population between 10 and 50 miles is expected to grow by 92 percent between 1990 and 2030, or from 659,411 to 1,266,338. The average annual growth rate for this area would be 2.3 percent for the 40 year period. This rate of growth can be compared to the rate for the State of Florida, which is expected to be 1.41 percent per year from 1990 to 2020. Florida is presently one of the most rapidly growing states in the United States. Between 1980 and 1990, the state grew by 32.7 percent, a net addition of over three million people. Nearly 87 percent of this growth was attributed to net migration.

Areas of development

The principal area of development between 10 and 50 miles of St. Lucie Plant occurs in Palm Beach County in the sectors including and adjacent to the Atlantic Coast. Major development activity outside of Palm Beach is concentrated in what can be called the "Atlantic Corridor", the 5 to 10-mile area between the Atlantic Ocean and either Interstate 95 or the Florida Turnpike in Martin, St. Lucie, Indian River, and southern Brevard Counties.

Land to the west of this region is mostly used for pasture, agricultural production (citrus, sugar cane, and truck farming), or remains undeveloped. Access is limited and population is sparse. In a few widely scattered sites, tracts of land have been platted and sold as home sites or proposed for such development. No significant development of any of these projects which lie west of the Atlantic corridor has yet taken place.

Development is focused in the Atlantic Corridor for the following reasons:

1. proximity to existing population centers and services;
2. access to the Atlantic Ocean and Indian River, and the amenities they provide: scenic beauty, sports and recreation, tourist and industry potential;
3. presence of soils suitable for development of the coastal ridge;
4. zoning and planning policies developed by county and regional agencies which permit development in these areas; and
5. availability of land suitable for development.

Only three significant clusters of development occur outside the Atlantic Corridor between 10 and 50 miles of St. Lucie Plant. Two are on or near the shores of Lake Okeechobee (which covers 400 square miles in sectors SW and WSW between 30 and 50 miles of the plant). On the southeastern shore of the lake in Palm Beach County, the community of Pahokee serves the agricultural community of the western section as well as the sport fishing community using the lake. A few miles north of the lake in Okeechobee County, a regional center has developed at Okeechobee City. The third location where significant development is occurring is Indiantown, in south central Martin County, at the intersection of the St. Lucie Canal and the Seaboard Coast Rail Line.

The following is a summary of development trends by county within 50 miles of St. Lucie Plant.

Palm Beach County. The principal area of growth within 50 miles of St. Lucie Plant is in the northeastern quadrant of Palm Beach County, which lies south of the plant, at a distance of more than 27 miles. About 40 percent of Palm Beach County falls within 50 miles of St. Lucie Plant; the total population of this area is expected to increase from 376,578 in 1990 to 733,136 in 2030. This increase represents a growth of 94.7 percent over the entire period, or 2.37 percent averaged annually. The corridor in Palm Beach County between the Atlantic Ocean and the Florida Turnpike is intensively developed with contiguous towns and cities such as Palm Beach, West Palm Beach, Riviera Beach, and Lake Park. Residential development activity is expected to continue in this area because of strong growth to date and its reputation as a desirable place to live. Many developments include self-contained recreation amenities. The Professional Golfer's Association (PGA) is headquartered in Palm Beach County. Another area of growth exists in the northwestern quadrant of Palm Beach County where Pahokee is located on the shore of Lake Okeechobee. Pahokee is the 15th largest city within 50 miles of St. Lucie Plant. It has an estimated population for 1990 of 6,822.

Martin County. Nearly all (approximately 84 percent) of Martin County's 1990 population is located between 10 and 50 miles of the plant. The remaining residents are within 10 miles of St. Lucie Plant. The 1990 total of 84,560 persons between 10 and 50 miles for this county is expected to grow by 119.7 percent to 101,218 by the year 2030. This represents an average annual growth rate of 2.99 percent. the city of Stuart is the major population center for the county; in 1990, its population of 11,936 represented 11.8 percent of the total county population of 100,900. Population is expected to grow in and around the city of Stuart and on the barrier beaches in the Atlantic Corridor in Martin County. Multi-family home construction decreased dramatically

between 1988 and 1989, when housing permits dropped by about a third. The decline has continued through the first half of 1992. Single family housing permits plummeted by nearly 40 percent across the Treasure Coast the from 1989 to 1990. Martin County has seen a slight increase in these permits in the first half of 1992 compared to 1991.

Indiantown is an incorporated area located approximately 26 miles southwest (immediately east of the dividing line between SSW and SW) of St. Lucie Plant at the intersection of SR 710 and SR 76. No population number has been published for Indiantown based on the 1990 census. The western part of Martin County is largely range and cropland, with few permanent residents outside of Indiantown.

St. Lucie County. St. Lucie County extends from the plant site west to the 30-mile radius. Of the county's total estimated population of 150,171 in 1990, approximately 22.3 percent, or 33,553 persons, are estimated to reside outside of the 10-mile radius. The number of people outside the 10-mile radius is expected to grow by 134.2 percent (or 3.35 percent average annual rate) to a population of 78,576 in 2030. The primary reason for this growth is the city of Port St. Lucie.

St. Lucie County's major population centers are the city of Fort Pierce, with a 1990 population of 36,830, and Port St. Lucie with a 1990 population of 55,866. While the county as a whole grew 72.3 percent between 1980 and 1990, the city of Fort Pierce grew 9.0 percent. Most of Fort Pierce's growth occurred in the portion of the city on Hutchinson Island. Inland, relatively little population growth is taking place in Fort Pierce. As Fort Pierce is built up, development is expected to occur within the Atlantic Corridor, outside the city limits. Growth in Port St. Lucie has occurred at a much faster pace than Fort Pierce. In 1980, Port St. Lucie's population was only 14,690. Since 1980, Port St. Lucie's population has increased by 280.3 percent.

Immediately to the west of the Atlantic Coastal Ridge is a freshwater marsh system known as the Savannas. Once found along the entire length of the Indian River Lagoon, this vanishing natural feature has been depleted by mans development. Through the continued effort of the State of Florida's Conservation and Recreational Land (CARL) acquisition program, privately held properties within this area are being acquired for perpetual public preservation. In addition to its inland estuary and isolated wetland network, St. Lucie County has 18 miles of Atlantic Ocean shoreline, much of which is currently undeveloped. Approximately 4.5 miles of this unincorporated oceanfront are under public ownership. FPL owns another 2 miles of oceanfront, and maintains the property in its natural state in conjunction with operation of the plant. The balance of the oceanfront properties is held in private ownership and is available for development activities, which have historically been residential in nature. About 40 percent of that property has already been developed.

The major use of land within the unincorporated areas of the county is agriculture. Well over 60 percent of the county is presently used for the production of citrus, cash crops, or ranching activities. The western portion of St. Lucie County is especially dominated by pasture and croplands. However, this situation should change dramatically by 2030, as the city of Port St. Lucie expands to the southwest. The locus of population in St. Lucie County will shift to this area. The largest urban use of land within the unincorporated area of the county is for detached, single family residential housing units. Multi-family and mobile home development activities account for a little less than a third as much acreage as the single family units. As was observed in Martin County, multi-family home construction decreased from 1988 to 1989, but even more drastically at a reduction of about two-

thirds. The decline has continued into 1992. The Treasure Coast trend of dropping single family housing permits is evident in St. Lucie County, with permits leveling off over 1991 through 1992.

Indian River County. Essentially all of Indian River County falls within the 10- to 50-mile radius. The county population of 90,208 in 1990 is expected to grow to 173,184 by 2030. This overall growth of 92 percent represents an annual average growth of 2.3 percent. The principal community in the Atlantic Corridor is the county seat, Vero Beach, with a 1990 population of 17,350 persons (19.2 percent of the total county population). Other cities and towns include Sebastian, 10,205 persons in 1990; and Indian River Shores, 2,278 persons in 1990. Only one settlement, the town of Fellsmere, with a 1990 population of 2,179, is located outside the Atlantic Corridor. Aside from the community at Fellsmere (NW 30-40), the area west of Interstate 95 is for the most part protected wetlands which are part of the St. Johns River Flood Control District.

Brevard County. The portion of Brevard County which lies within the 50-mile radius of St. Lucie Plant is included in sector-segments NNW 30-40, NNW 40-50, and NW 40-50. From the 1980 decennial census to 1990, Brevard tied with Okeechobee County for sixth place in percent growth (46.2 %) of the nine counties involved in the 50-mile radius. Only the inland counties of Highlands and Glades experienced slower growth, thus Brevard had the slowest growth rate of the coastal counties. Major development in Brevard County over 1980 to 1990 has taken place at Cape Canaveral, Malabar, Palm Shores, Palm Bay, and West Melbourne. It is expected that the main growth will continue in the southern portion of the county as vacant coastal areas become more attractive.

In southern Brevard County, development has occurred along the Indian River and Atlantic Coast. Small communities include Barefoot

Bay, Micco, Grant, Valkaria, Melbourne Shores, and Floridana Beach. The only incorporated town entirely within the 50-mile radius of St. Lucie Plant is Malabar, which in 1990 had a population of 1,977. The town of Palm Bay lies to the north of Malabar, just outside the 50-mile radius, on the Indian River. However, part of Palm Bay's incorporated area falls within the 50-mile radius in NNW 40-50 and NW 40-50. The development of Palm Bay is proceeding at a rapid pace. The city grew from 18,560 persons in 1980 to 62,632 in 1990, an increase of 237.5 percent. In southern Brevard, as in Indian River County, development will be confined to the eastern coastal area because of restrictions imposed in the western region by the St. Johns River Flood Control District.

Okeechobee County. Most of Okeechobee County's 1990 population of 29,627 persons resides within the 10- to 50-mile area. By the year 2030, this number is expected to increase by 96.5 percent to 58,219. Okeechobee's population is concentrated in and around the county seat of Okeechobee City. The county seat is at the convergence of US 98 and US 441 and SR 70, SR 78, and SR 710, less than 5 miles north of Lake Okeechobee. This accessibility is expected to ensure it's continued growth as a regional center. The city's 1990 population of 4,943 represents about 16.7 percent of the county total. The unincorporated population of 24,684 persons accounts for the other 83.3 percent of the county's total population. Along with Okeechobee City, the nearby community of Cypress Quarters is split between sector-segments WSW 30-40 and W 30-40. In addition, WSW 30-40 contains Taylor Creek. W 30-40 encompasses Basswood Estates, Whispering Pines, and Country Hill Estates.

Glades, Osceola, and Highlands Counties. Three counties on the periphery of the 50-mile study area contribute less than half of a

percent of the 1990 population between 10 and 50 miles of St. Lucie Plant. In Glades County, in sector-segment WSW 40-50 , on the northwest shore of Lake Okeechobee, a community known as Buckhead Ridge has developed. The only other settlement of greater size in Glades County is the county seat of Moore Haven, which had a 1990 population of 1,432, an increase of 14.6 percent from 1980. The unincorporated portion of Glades County accounted for 6,159 persons in 1990, representing an increase of 29.9 percent over 1980.

Osceola County is partially included in the 50-mile radius in sectors NW and WNW. The only significant settlement is Yeehaw Junction, found in sector-segment WNW 40-50.

Like Osceola, Highlands County has a small fraction of its land area within the 50-mile radius, In this area, a small settlement has developed on SR 70. Highlands County's predominant growth is expected to continue outside of the 50-mile radius in the vicinity of Sebring, Avon Park and Lake Placid, in the central part of the county. All three interior counties reflect the lower rates of development taking place in Florida's central regions, which are not adjacent to the Atlantic or Gulf coasts.

Projected Growth Rates Between 10 and 50 Miles

The total population between 10 and 50 miles is expected to grow by 92 percent from an estimated 659,411 persons in 1990 to 1,266,338 in 2030. Regarding the six main counties involved in the 50-mile radius, the projections listed in Table 3-2 are expected. Sector-segments likely to have the highest rates of growth are those that include attractive areas near population centers that have reached, or are approaching saturation. The listing above indicates a growth hierarchy for the top three counties of St. Lucie > Martin > Palm Beach.

Table 3-2 Projections to 2030 by County

| County | 1990 Pop. | 2030 Pop. | Tot. Change | % Inc. |
|---|---|---|---|---|
| Brevard | 398,978 | 779,646 | 380,668 | 95.4 |
| Indian River | 90,208 | 173,184 | 82,976 | 92.0 |
| Martin | 100,900 | 221,676 | 120,776 | 119.7 |
| Okeechobee | 29,627 | 58,219 | 28,592 | 96.5 |
| Palm Beach | 863,518 | 1,750,700 | 887,182 | 102.7 |
| St. Lucie | 150,171 | 351,678 | 201,507 | 134.2 |

An analysis of 1990 resident populations per square mile of land area available shows that sector-segment SSE 40-50 has the highest density of 3,188 persons per square mile. This area is entirely within Palm Beach County and includes the city of West Palm Beach. Expansion is likely directly north into SSE 30-40; also dense, but to a lesser degree at 1,594 persons per square mile. Growth could also continue to the west in S 40-50 where the current density is 867 persons per square mile. The aforementioned sector-segments of likely expansion are also entirely within Palm Beach County. In contrast, a relatively low rate of growth is expected for annular sector SSE 40-50 because of its high density. Low growth rates can also be expected for sector-segments of low density surrounded by others of low density (less than 500 persons per square mile). Such sector-segments include SW 20-30, WSW 20-30, W 20-30, WNW 20-30, W 30-40, and WNW 30-40.

The focus of this chapter has been the estimation of populations in the recent past but after the latest decennial census, and the projection of populations into the distant future. Conversely, the next chapter provides techniques to reconstruct populations bounded by past decennial censuses, but at smaller geographic levels than originally reported.

CHAPTER 4
DEVELOPMENT OF POPULATION RECONSTRUCTIONS

## Overview

Data from past national decennial censuses along with county data are plotted to yield growth curves (usually exponential) for regions of interest. The equation describing the curve is solved for unknown population values at some past time by assuming that subpopulations within a region of known growth also grew at a proportional rate. It is assumed that the uncertainty increases as the absolute value of the difference in the present and the reconstructed time frame increases (as with projections), however important components may be largely subjective. Recent data with better measures of counting error are quantifiable. Data available on CD-ROM at the block group level facilitates a modified housing unit method where modern housing unit densities can be correlated with their average ages to estimate inhabitance during past years.

However, of particular interest are techniques based specifically on population dynamics, e.g., births, deaths, migration, as opposed to secondary indicators like housing units. Retrospective analyses have advantages over prospective ones since limits are available for each decennial census at the level of its smallest (or most convenient) areal reporting. Limiting brackets consisting of successively enumerated decennial censuses aid intercensal estimations. Methodological strategies to estimate intercensal populations are numerous. These

strategies vary in complexity, input data requirements, and level of detail obtained in the final computation. Some methods allow considerable flexibility in the way they are used. Regardless of the particular strategy, they all are based on the classic demographic equation that relates the population at some point in time to the population at a previous point in time by natural increase (births minus deaths), and net migration. Annual intercensal estimations are often made by some type of interpolative process. The beginning and ending points are two consecutive decennial census enumerations, and an assumption is made that population change is continuous and smooth. There are three major techniques based on interpolation for estimation of populations which differ in their treatment of the nature of population growth. These are a simple arithmetic rate of change, a compound rate of growth expressed as a geometric rate of change, or a compound rate expressed as an exponential rate of change. The geometric rate assumes a compound rate of growth over a specified time period, e.g., annually, and the exponential rate asssumes a continuous compound rate of growth. In a time span of 10 years, growth rates determined by geometric and exponential means are quite similar. Furthermore, results obtained using arithmetic versus compound growth rates will not differ greatly over such a short time span. There is no clear advantage between the strategies, but a compound growth rate is usually chosen to agree with the model for human reproduction. An advantage that all of these interpolation techniques share is that they may be applied to any level of geography from the entire world to the census block level. In a detailed GIS procedure, the smallest common enumeration boundaries between consecutive censuses, e.g., block groups or tracts, can be analyzed and estimated independently and then aggregated into sector-segments. In reality, the assumption of continuous and constant change

is erroneous in that a small population may sporadically increase, decrease, or stagnate. At the levels of geography of concern in this study, i.e., sub-state and county level, net migration is an important factor in population change and is likely to be sporadic. Also, if a researcher desires detailed information on race, age, or sex, the interpolation approach is not a good choice. When such detail is needed, the forward/reverse survival rate method is useful. For this method, the data requirements are more rigorous and the calculations are more difficult than the interpolation methods. Forward survival rates are applied to the base population enumerated in the first census, and the results are then averaged with those obtained by application of the reverse survival rates to the subsequent census. Appropriate detailed survival rates are less likely to be available than are detailed population data, especially at smaller levels of geography. This is why it is advisable to check the availability and detail of the survival rates data before deciding the level of estimation detail to be calculated. All things being equal, a population estimate for a given area for a given point in time is preferable to a population projection for the same area and time period.

The greater the level of detail in the population estimates, the greater the level of error. Population estimates of the total population are generally more accurate than estimates that include some form of population detail such as age, race, or sex. Detailed population estimates require a greater level of data and more complex methodology than estimates for population totals only.

Estimates produced through the use of direct demographic data will be more accurate than estimates produced through the use of indirect or symptomatic data. Direct data are those involving any of the time parameters of population change: fertility, mortality, and migration.

Potential indirect data sources would include a large number of variables such as school enrollments, automobile registrations, voting records, utility service hookups, and a variety of others. If given the choice, no one would choose to employ indirect data over direct data, but indirect data are often the only data available. The migration component of population change is especially likely to be estimated only through indirect data, particularly when population estimates are made at subnational levels of geography.

Beyond the issue of direct versus indirect data, the higher the quality of input data, the higher the quality of the resulting estimates. Most estimation methodologies begin with a base population at a previous point in time. A base population consisting of the most recent census is of higher quality and will yield better current estimates than a base population consisting of a previous estimate. A similar argument can be made for other data sources. Data resulting from official sources whose collection is mandatory will almost always be of higher quality than data collected on an informal basis.

The longer the period of time between the enumeration of the base population and the estimation date, the greater the degree of error in the estimate. Estimates made for shorter periods of time--for example, for 2 to 3 years past the base period--will generally be more accurate than estimates made 8 or 9 years past the base year. Ironically, the need for current estimates is more critical the farther one moves from the base period.

In the most general case, the simpler mathematical methods of population estimation will produce results with greater error than the more complex methods of estimation. Simpler mathematical methods of population estimation include interpolation, extrapolation, and proration. Interpolation techniques are most often used for intercensal

estimates where the population is known at a beginning point and an end point but not for the intervening years. Extrapolation methods are essentially a simple type of population projection technique, and proration methods involve allocating the known population of a larger area among the several smaller subareas comprising the larger area. More complex methods of population estimation will include cohort techniques where each major age group of the population will be estimated separately, ratio correlation techniques which involve the development of a statistical regression equation, and several others.

The mathematical methods will usually result in estimates of poorer quality than the more complex methods, with mathematical extrapolation techniques resulting in the least accurate estimates. It should be noted, however, that the use of a more complex method does not necessarily guarantee a quality result. The more complex the method, the more demanding the data requirements, and in some cases the required data will simply not be available or will be of poor quality. No method of population estimation based on poor quality data will produce an accurate result, and complex methods of estimation are no exception. In fact, in the absence of quality data, a simpler method of population estimation may be the best choice.

No single method of population estimation will always be the best choice. Over time, one will achieve better results from a program of population estimation by employing a variety of estimation techniques. Using multiple techniques will provide a means of checking the validity of the estimate since similar results obtained from a variety of different methods tend to suggest the overall accuracy of the result. In many cases, one will be well served by averaging the results of several estimates to obtain the final population estimate.

## Methods of Intercensal Population Estimation

The methods of population estimation are employed in two broad applications: intercensal estimation and postcensal estimation. Intercensal estimates result when one has population totals for a particular area from two successive censuses for the census years but does not have population totals for the intervening years. While one may deal with an intercensal period or a postcensal period, in each case one is dealing with estimates in contrast to projections. Recall that the essential difference between a population estimate and a population projection is that population estimates are made for an area for a specific time in the past. Most often the estimate is for a period in the recent past, such as the previous year; population projections are produced for a period of time in the future. In most cases the methodologies from which one will choose when producing population estimates or projections will differ due to the applicable time frame. In addition, methods of population estimation are more likely to be based on the use of symptomatic data which are likely to move in conjunction or sympathy with population totals. Methods of population projection are more likely to be based on some type of extension or variation of current population trends.

A number of methodological strategies to estimate intercensal population have been developed. The methods vary in their level of complexity, type of data required for their computation, level of detail in the final product, and intended purpose. In addition, many of the methods of producing intercensal population estimates are better described as methodological strategies; that is, there is considerable flexibility in the way certain estimation methods are employed. One common characteristic of population estimation methodologies is that

they all in one way or another are based on the classic demographic equation. Not all of the estimation methodologies, however, will deal with the classic demographic equation in as sophisticated a fashion as others.

Not all of the existing methodologies are described in the present discussion. Rather, emphasis is placed on those methods that are used more frequently by those active in population estimation. Included in this discussion of estimation methodologies are intercensal interpolation, assuming arithmetic and compound rates of change, and detailed intercensal estimates through the forward/reverse survival rate method. The discussion of intercensal estimation methodologies begins with an analysis of the interpolation method of producing intercensal estimates.

Interpolation in Intercensal Estimates

Intercensal estimation most often involves the production of single year population estimates for all years between two census years. For example, we may have the 1980 and 1990 census population for an area but would like to have a year-by-year estimate for years 1981 through 1989. Often some type of interpolation estimation process will be chosen to derive the intercensal estimates. The two observed population totals reported in the two most recent census results are treated as the beginning and ending points for the population, and the interpolation procedure then provides a series of annual estimates. One should note that implicit in an interpolation strategy is the assumption that the population has increased or decreased in a smooth and continuous fashion between the two census years. In many cases the assumption of a continuous population change will be false. Population change often occurs in an irregular fashion with periods of rapid change followed by

periods of stagnation. In the case of intercensal estimates, however, the fact that the beginning and end points of the annual population series are represented by actual census results tends to lessen the importance of the estimated population values for the intervening years.

There are three major variations of producing population estimates through interpolation. The chief difference among the three variations is the assumption concerning the nature of population growth that has occurred during the period between the two most recent censuses. One may assume a simple rate of population growth between the two most recent census periods as expressed by an arithmetic rate of change, or one may assume a compound rate of growth between the two most recent census periods. The compound rate of growth may be expressed as a geometric rate of change or as an exponential rate of change. A geometric rate of change assumes a compound rate of growth over a specified finite time period, such as compounding annually, quarterly, weekly, and so on. An exponential rate of change assumes a continuous compound rate of growth with the period of compounding infinitely short.

Arithmetic method

An arithmetic rate of change is the simplest approach in producing intercensal population estimates. One begins with the observed population at each of two census years; for example, 1980 and 1990. The population for the earlier period is subtracted from the later period to determine the amount of net change in population between the two census years. The net change in population is then divided by the number of years in the time interval between the two censuses to determine the annual amount of population change between the two time periods. The annual amount of population change is then added to the earlier census total to derive the estimated population for the first year after the census. The annual amount of population change is then added to the

previous year's estimate to derive the second year estimate, and so on until all of the intercensal years have been estimated. In the following example, it is assumed that one has population data for the 1980 and 1990 census years.

$$\textit{Annual Net Change} = \frac{(P90 - P80)}{n}$$

where P90 is the 1990 census population
P80 is the 1980 census population
n is the number of years between the two census periods

For Pinellas County
Population 1980: 728,531
Population 1990: 851,659
Annual Net Change = (851,659 - 728,531)/10 = 12312.8

The series of intercensal estimates would be derived by adding 910 persons to the preceding population total on a year-by-year basis as follows.

Table 4-1 Intercensal Estimates

| Year | Population |
|---|---|
| 1980 | 728,531 |
| 1981 | 740,844 |
| 1982 | 753,157 |
| 1983 | 765,470 |
| 1984 | 777,782 |
| 1985 | 790,095 |
| 1986 | 802,408 |
| 1987 | 814,721 |
| 1988 | 827,033 |
| 1989 | 839,346 |
| 1990 | 851,659 |

Population totals for 1980 and 1990 represent census values for Pinellas County, and the population values for years 1981 through 1989 represent estimates based on an arithmetic rate of growth.

Census results are reported as of April 1 in the United States. In many applications the convention for producing population estimates is to use a July 1 date for the estimate years in order to differentiate more clearly a population estimate from an actual census result. When a July 1 date is desired for a series of population estimates, it is necessary to adjust the series since the first estimate occurs 1.25 years past the census date. In order to adjust the above series, one would multiply the annual amount of population increase (12312.8 persons) by 1.25 years (the period elapsed since the base population was enumerated and the first estimate is produced). In the present example, the result of the multiplication is 15,391 persons. The estimate for 1981 would then be calculated by adding 15,391 to the base population of 728,531 for an estimate of 743,922. The remainder of the series would be derived by adding the annual amount of net change (12312.8 persons) to the estimated population of the preceding year. The final population value in the series represented by the census count of 1990 will be 0.75 x 12312.8 (persons) above the 1989 estimate since April 1, 1990, is only three-quarters of a year past July 1, 1989.

In the previous example, the annual net change in population was calculated by subtracting the beginning 1980 population from the ending 1990 population, in order to derive the total net change, and then dividing the total net change by the number of years between the two census periods. The result was an annual net change of 12312.8 persons per year. One may choose to divide the annual net change by the size of the base population (the 1980 population of 728,531) and then multiply by 100 to determine the annual percentage change in population. In this

example, the annual percent change would be (12312.8/728,531) x 100 = 1.69%.

Compound methods

There are two procedures for producing intercensal estimates assuming compound rate of change. One method assumes a geometric rate of change which involves compound growth over a finite period of time. Usually, one year is chosen as the period of time for compounding. The second method of intercensal estimation assumes an exponential compound rate of change. With an exponential rate of change, the compound growth is assumed to occur on a continuous basis. Rather than having a finite period of time such as a year, half-year, quarter, month, or week over which to calculate the growth rate, an exponential method assumes an infinitely small period of time.

Geometric growth

The geometric rate of change assumes the population is increasing by a constant amount compounded over a given period of time. In the general case, the formula for a geometric rate of population growth is as follows:

$$P_n = P_0 \cdot (1 + r)^n$$

where $P_n$ is the end point population
$P_0$ is the initial population
$1 + r$ is the annual rate of growth
$n$ is the number of years between the base population and the estimate period

Assuming census population data for 1980 and 1990, a formula for a geometric rate of growth is as follows:

$$P_{90} = P_{80} \cdot (1 + r)^n$$

where $P_{90}$ is the 1990 census population
$P_{80}$ is the 1980 census population

$1 + r$ is the annual rate of growth
$n$ is the number of years between the base population and the estimate period

The rate of population growth $(1 + r)$ may be calculated by taking the $n^{th}$ root of $P_n/P_0$. For example, if there were 2 years between the 2 population observations, then one would take the square root of the result of $P_n/P_0$. In the present case, there are 10 years between the 2 census observations so one would take the $10^{th}$ root of $P_{90}/P_{80}$. Readers should note that there are alternate methods for calculating the value of $(1 + r)$ involving logarithmic transformations.

Assume the same 1980 and 1990 census data as in the previous example where Population 1980 is 728,531 and Population 1990 is 851,659. Then $P_{90}/P_{80} = 1.169$; solving for the $10^{th}$ root of 1.169, the result is 1.01574. The result of 1.01574 represents the value of $(1 + r)$. The value of r (0.01574) represents the proportion of compound annual population growth. Expressed as a percentage, the value would be 1.574% which may be compared to the annual percentage growth rate of 1.69% found in the previous example where an arithmetic rate of population growth was assumed.

Population estimates for the intercensal years are obtained by substituting the appropriate values into the formula. The estimate for 1981 is obtained as follows: $P_{1981} = 728,531 \times (1.01574)^1 = 739,998$. The estimate for 1982 is obtained as follows: $P_{1982} = 728,531 \times (1.01574)^2 = 751,646$.

The population estimates assuming a geometric rate of growth may be adjusted to a July 1 date just as the arithmetic rate of growth estimates were. In the case of the geometric rate estimates, one would simply substitute 1.25 as the exponent n in the term $(1 + r)^n$ rather than an exponent of 1 used in the example above for the first year's

estimate. The second year's estimate would use the exponent 2.25 and so on throughout the series.

Exponential growth

The exponential rate of growth is the second intercensal estimate strategy based on a compound rate of growth. The exponential rate of growth assumes a continuous period of compounding rather than compounding over a finite period, such as annually, monthly, weekly, and so on. In some ways the exponential rate of growth is quite similar to what we have seen with the geometric rate of growth, but the calculations may seem a little more cumbersome.

Exponential rates of growth may be expressed as a function of e, the natural logarithm whose value is approximately 2.718. The population at a given time may be expressed as the population at the base period multiplied by e raised to a power reflecting the assumed rate of growth times the time interval since the base period. Using the same census data for 1980 and 1990 as above, the 1990 population may be expressed as follows:

$$P_{90} = P_{80} \cdot e^{rn}$$

where
- $P_{90}$ is the 1990 census population
- $P_{80}$ is the 1980 census population
- e is the natural logarithm
- r is the growth rate
- n is the number of years between the base population and the estimate period

The census data are given, and n is equal to 10 years. In order to complete the calculations, it is necessary to find the value of r, the growth rate between the two census periods. The calculation for r is simply a matter of substituting the appropriate values into the following formula:

$$r = \frac{\log \frac{P_{90}}{P_{80}}}{n \cdot \log e}$$

To solve the formula for r, we simply divide the 1990 population by the 1980 population and then take the base 10 logarithm of the result. The resulting term is then divided by the product of the number of years between the two population observations (in this case n = 10 years) and the base 10 logarithm of e. Substituting the appropriate values into the formula yields:

r = (log (851,659/728,531)) / (10 x log e)
= (log(1.169)) / (10 x 0.43429)
= 0.01561

The value r = 0.01561 represents the exponential growth rate of the population between 1980 and 1990. Expressed as a percentage, the growth rate is 1.56% compounded continuously over the 10-year period. The exponential growth rate may be compared to the geometric growth rate of 1.57% compounded annually over the 10-year period and the arithmetic growth rate of 1.69%.

Once the value of r has been calculated, the intercensal estimates may be computed on a year-by-year basis by substituting the appropriate value of n, representing the time between the base population year and the estimate year, into the estimating formula. The population estimate for 1981 would be obtained as follows:

$P_{1981}$ = 728,531 x $2.718^{(0.01561)(1)}$
= 728,531 x 1.01573
= 739,993.

The estimates for the remaining years in the series would be obtained by substituting the appropriate values of n into the formula. For example, the 1982 estimate would be obtained by using n = 2, etc.

The exponential growth rate method is also easily adapted to a July 1 date for the annual estimate period. One would simply adjust the value of n indicating the time of the estimate period beyond the base population. A July 1, 1981, estimate would be calculated by using a value of n = 1.25 to reflect an estimate date 1.25 years past the time of the base population.

Comparing Results from Arithmetic and Compound Rates of Growth

Since the previous examples deal with intercensal estimates there is no difference between arithmetic and compound rates of growth with respect to the beginning point and the end point in the estimation process. All three types of intercensal estimation strategies examined begin with the population enumerated in the second to most recent census and end with the population enumerated in the most recent census. The results obtained for the simple arithmetic growth rate assumption differ compared to the results from the compound growth rate assumption, but the 10-year interval between the two census observations does not provide enough time for major differences to emerge.

The intercensal estimates obtained with the arithmetic growth rate method change by a fixed absolute amount from year to year but by a variable percentage from year to year. The intercensal estimates obtained with the compound growth rate assumptions change by a variable absolute amount from year to year but by a fixed percentage from year to year. Given the small degree of difference between the two sets of results, there is no real advantage or disadvantage to either strategy. Conventional wisdom suggests that a strategy assuming a compound growth rate is more realistic since human reproduction follows that model, but as we have seen in the case of intercensal estimates the beginning and end points of the series are fixed to equal the observed census results

of the two most recent censuses, and there is not that much time for the results of the two strategies to vary that much. There is also the implicit assumption of a relatively smooth rate of change involving gradual growth, or in some cases, gradual decline from beginning point to end point. In reality, intercensal population change is much more sporadic. Periods of growth or decline may be followed by an opposite trend or by a period of stagnation. In such cases, neither method will provide an accurate picture of population change between the two census counts. When working with small levels of geography, such as a state or county, the role of migration becomes increasingly important with respect to population change. Population growth due to net migration is not necessarily better described by a compound growth rate assumption, and, in fact, growth due to migration is likely to occur in sporadic periods following perceived economic opportunities in the area.

## Key Issues with Interpolation Techniques

The three types of interpolation techniques used to produce intercensal estimates--arithmetic growth assumption, geometric growth assumption, and exponential growth assumption--are relatively easy to calculate, and they have a wide range of application. As illustrated previously, when these methods are used for intercensal estimates, there is not a great deal of difference in the results, so one approach does not have a clear advantage over another. All three of the techniques are versatile in that they may be applied to multiple levels of geography. Even the smallest unit of geography, such as a census tract or block, may be estimated using interpolation procedures. All that is required are beginning and ending census results.

One limitation of an interpolation methodology is that it does not result in detailed population estimates. The interpolation techniques are best suited to the production of total population estimates rather than estimates by race, age, and sex. The interpolation approach is also limited by what are often unrealistic assumptions on the nature of population change between census periods. All three of the strategies assume a rather smooth rate of change from one census period to the next which is not often the way populations change. In spite of some limitations, the interpolation approach to producing intercensal estimates is an efficient one.

Each of the interpolation techniques may be used for the development of postcensal estimates. Postcensal estimates are made for years following a census but differ from the intercensal estimates in that they are produced without reference to an end point census value. In effect, the use of interpolation techniques for postcensal estimates results in more of a population projection than a true population estimate.

## Case Study: Pinellas Plant Population Reconstruction

Reconstucted populations by sector-segment were calculated for decennial census years 1960, 1970, and 1980 by using the new GIS-based proration technique described in Chapter 3 for projections. County census counts were used to find growth ratios between consecutive censuses as shown in Table 4-2. These ratios were applied in the DBMS to census blocks from those respective counties, and the population assignment to the new regions (sector-segments) was performed. The resulting decennial reconstructions are submitted in Tables 4-3 through 4-8. It is a simple matter to apply any of the mathematical intercensal

interpolation strategies discussed in this chapter to these prorated and area-weighted populations to yield estimates for any date between those years.

Table 4-2 County Growth Ratios

| | Population by Census Year | | | | Censal Ratio | | |
|---|---|---|---|---|---|---|---|
| COUNTY | 1990 | 1980 | 1970 | 1960 | 80/90 | 80/70 | 70/60 |
| Hardee | 19,499 | 20,357 | 14,889 | 12,370 | 1.0440 | 0.7314 | 0.8308 |
| Hernando | 101,115 | 44,469 | 17,004 | 11,205 | 0.4398 | 0.3824 | 0.6590 |
| Hillsborough | 834,054 | 646,939 | 490,265 | 397,788 | 0.7757 | 0.7578 | 0.8114 |
| Manatee | 211,707 | 148,445 | 97,115 | 69,168 | 0.7012 | 0.6542 | 0.7122 |
| Pasco | 281,131 | 193,661 | 75,955 | 36,785 | 0.6889 | 0.3922 | 0.4843 |
| Pinellas | 851,659 | 728,531 | 522,329 | 374,665 | 0.8554 | 0.7170 | 0.7173 |
| Polk | 405,382 | 321,652 | 228,515 | 195,139 | 0.7935 | 0.7104 | 0.8539 |
| Sarasota | 277,776 | 202,251 | 120,413 | 76,895 | 0.7281 | 0.5954 | 0.6386 |

Table 4-3 1980 Reconstructed Populations by Sector-Segment to 10 Miles

| 1980 RECONSTRUCTED POPULATIONS BY SECTOR-SEGMENT | | | | | |
|---|---|---|---|---|---|
| | DISTANCE (MILES) | | | | |
| DIRECTION | 0-2 | 2-3 | 3-4 | 4-5 | 5-10 |
| N | 2,071 | 4,986 | 4,760 | 5,785 | 48,333 |
| NNE | 3,256 | 3,617 | 3,759 | 3,525 | 21,511 |
| NE | 2,219 | 1,529 | 3,664 | 1,306 | 4 |
| ENE | 1,531 | 512 | 766 | 1,044 | 1,212 |
| E | 2,188 | 979 | 711 | 420 | 10,742 |
| ESE | 1,625 | 4,674 | 5,937 | 3,807 | 41,867 |
| SE | 1,555 | 3,014 | 4,885 | 4,774 | 60,343 |
| SSE | 1,366 | 3,813 | 5,639 | 7,653 | 55,869 |
| S | 1,988 | 1,299 | 4,985 | 5,029 | 26,337 |
| SSW | 3,693 | 2,712 | 2,119 | 4,032 | 5,586 |
| SW | 1,957 | 1,361 | 4,920 | 4,480 | 4,014 |
| WSW | 1,347 | 2,923 | 5,700 | 5,698 | 5,281 |
| W | 2,028 | 4,409 | 2,598 | 5,979 | 7,743 |
| WNW | 1,831 | 4,275 | 4,600 | 7,482 | 6,022 |
| NW | 1,328 | 3,007 | 4,869 | 7,071 | 7,229 |
| NNW | 867 | 2,948 | 5,086 | 6,753 | 35,700 |

Table 4-4 1980 Reconstructed Populations by Sector-Segment to 50 Miles

| 1980 RECONSTRUCTED POPULATIONS BY SECTOR-SEGMENT | | | | |
|---|---|---|---|---|
| | DISTANCE (MILES) | | | |
| DIRECTION | 10-20 | 20-30 | 30-40 | 40-50 |
| N | 92,153 | 76,469 | 32,310 | 2,341 |
| NNE | 28,293 | 11,172 | 16,598 | 35,649 |
| NE | 87,921 | 84,727 | 14,083 | 35,238 |
| ENE | 108,941 | 143,978 | 39,174 | 52,770 |
| E | 41,036 | 64,701 | 27,411 | 39,907 |
| ESE | 1,964 | 25,706 | 2,135 | 427 |
| SE | 16,073 | 9,823 | 2,352 | 1,630 |
| SSE | 28,255 | 69,630 | 98,031 | 91,523 |
| S | 6,392 | 8,856 | 1,895 | 0 |
| SSW | 0 | 0 | 0 | 0 |
| SW | 0 | 0 | 0 | 0 |
| WSW | 0 | 0 | 0 | 0 |
| W | 0 | 0 | 0 | 0 |
| WNW | 0 | 0 | 0 | 0 |
| NW | 0 | 0 | 0 | 0 |
| NNW | 2,531 | 0 | 0 | 0 |

Table 4-5 1970 Reconstructed Populations by Sector-Segment to 10 Miles

| 1970 RECONSTRUCTED POPULATIONS BY SECTOR-SEGMENT | | | | | |
|---|---|---|---|---|---|
| | DISTANCE (MILES) | | | | |
| DIRECTION | 0-2 | 2-3 | 3-4 | 4-5 | 5-10 |
| N | 1,485 | 3,575 | 3,412 | 4,148 | 34,653 |
| NNE | 2,334 | 2,593 | 2,695 | 2,527 | 15,422 |
| NE | 1,591 | 1,096 | 2,627 | 937 | 3 |
| ENE | 1,098 | 367 | 550 | 748 | 869 |
| E | 1,569 | 702 | 510 | 301 | 7,701 |
| ESE | 1,165 | 3,351 | 4,256 | 2,730 | 30,017 |
| SE | 1,115 | 2,161 | 3,503 | 3,423 | 43,263 |
| SSE | 979 | 2,734 | 4,043 | 5,487 | 40,056 |
| S | 1,425 | 931 | 3,574 | 3,606 | 18,883 |
| SSW | 2,648 | 1,944 | 1,519 | 2,891 | 4,005 |
| SW | 1,403 | 976 | 3,527 | 3,212 | 2,878 |
| WSW | 966 | 2,096 | 4,086 | 4,085 | 3,787 |
| W | 1,454 | 3,161 | 1,863 | 4,286 | 5,552 |
| WNW | 1,312 | 3,065 | 3,298 | 5,364 | 4,318 |
| NW | 952 | 2,156 | 3,491 | 5,070 | 5,183 |
| NNW | 621 | 2,113 | 3,646 | 4,841 | 25,596 |

Table 4-6 1970 Reconstructed Populations by Sector-Segment to 50 Miles

| 1970 RECONSTRUCTED POPULATIONS BY SECTOR-SEGMENT | | | | |
|---|---|---|---|---|
| | DISTANCE (MILES) | | | |
| DIRECTION | 10-20 | 20-30 | 30-40 | 40-50 |
| N | 66,070 | 29,992 | 12,672 | 895 |
| NNE | 20,747 | 5,584 | 6,510 | 13,737 |
| NE | 66,629 | 54,915 | 6,296 | 13,803 |
| ENE | 82,558 | 109,110 | 28,971 | 33,327 |
| E | 31,098 | 49,032 | 20,773 | 28,824 |
| ESE | 1,488 | 19,348 | 1,540 | 298 |
| SE | 11,852 | 6,732 | 1,539 | 1,047 |
| SSE | 20,258 | 45,553 | 62,114 | 54,490 |
| S | 4,583 | 5,794 | 1,240 | 0 |
| SSW | 0 | 0 | 0 | 0 |
| SW | 0 | 0 | 0 | 0 |
| WSW | 0 | 0 | 0 | 0 |
| W | 0 | 0 | 0 | 0 |
| WNW | 0 | 0 | 0 | 0 |
| NW | 0 | 0 | 0 | 0 |
| NNW | 1,815 | 0 | 0 | 0 |

Table 4-7 1960 Reconstructed Populations by Sector-Segment to 10 Miles

| 1960 RECONSTRUCTED POPULATIONS BY SECTOR-SEGMENT | | | | | |
|---|---|---|---|---|---|
| | DISTANCE (MILES) | | | | |
| DIRECTION | 0-2 | 2-3 | 3-4 | 4-5 | 5-10 |
| N | 1,065 | 2,564 | 2,448 | 2,975 | 24,857 |
| NNE | 1,674 | 1,860 | 1,933 | 1,813 | 11,062 |
| NE | 1,141 | 786 | 1,884 | 672 | 2 |
| ENE | 787 | 264 | 394 | 537 | 623 |
| E | 1,125 | 504 | 366 | 216 | 5,524 |
| ESE | 836 | 2,404 | 3,053 | 1,958 | 21,531 |
| SE | 800 | 1,550 | 2,512 | 2,455 | 31,033 |
| SSE | 703 | 1,961 | 2,900 | 3,936 | 28,732 |
| S | 1,022 | 668 | 2,564 | 2,586 | 13,544 |
| SSW | 1,899 | 1,395 | 1,090 | 2,074 | 2,873 |
| SW | 1,007 | 700 | 2,530 | 2,304 | 2,064 |
| WSW | 693 | 1,503 | 2,931 | 2,930 | 2,716 |
| W | 1,043 | 2,267 | 1,336 | 3,075 | 3,982 |
| WNW | 941 | 2,198 | 2,365 | 3,848 | 3,097 |
| NW | 683 | 1,546 | 2,504 | 3,636 | 3,718 |
| NNW | 446 | 1,516 | 2,615 | 3,473 | 18,360 |

Table 4-8 1960 Reconstructed Populations by Sector-Segment to 50 Miles

| 1960 RECONSTRUCTED POPULATIONS BY SECTOR-SEGMENT | | | | |
|---|---|---|---|---|
| | DISTANCE (MILES) | | | |
| DIRECTION | 10-20 | 20-30 | 30-40 | 40-50 |
| N | 47,392 | 14,525 | 6,137 | 590 |
| NNE | 15,663 | 3,226 | 3,153 | 8,332 |
| NE | 54,061 | 39,168 | 3,358 | 6,806 |
| ENE | 66,985 | 88,529 | 23,033 | 24,267 |
| E | 25,232 | 39,783 | 16,854 | 24,308 |
| ESE | 1,208 | 15,602 | 1,196 | 235 |
| SE | 9,059 | 4,995 | 1,096 | 730 |
| SSE | 14,531 | 32,444 | 42,639 | 34,797 |
| S | 3,287 | 4,126 | 883 | 0 |
| SSW | 0 | 0 | 0 | 0 |
| SW | 0 | 0 | 0 | 0 |
| WSW | 0 | 0 | 0 | 0 |
| W | 0 | 0 | 0 | 0 |
| WNW | 0 | 0 | 0 | 0 |
| NW | 0 | 0 | 0 | 0 |
| NNW | 1,302 | 0 | 0 | 0 |

## Summary

Intercensal population estimates have a variety of uses and may be produced through a number of methodological strategies. Mathematical interpolatiòn techniques play a major role in the production of intercensal population estimates. One major use of interpolation techniques is the direct production of intercensal estimates. Given census results representing a beginning point and an end point for population size, interpolation may be used to produce an annual series of intercensal estimates. Rates of growth for the estimate series may be based on a simple arithmetic assumption or on a compound rate of growth reflecting either a geometric or exponential rate of change. One may continue to apply the assumed rates of growth beyond the most recent census as a way of producing a set of postcensal population estimates;

however, interpolation techniques used in a postcensal situation are described more accurately as population projections rather than as estimates.

The interpolation strategies for producing intercensal population estimates are quite versatile in that they may be applied to any level of geography. The data requirements are minimal, and the calculations are relatively easy. Interpolation methodologies are limited in that they generally produce only estimates of population totals, not detailed estimates. In addition, the use of interpolation in the production of intercensal population estimates assumes that the population change occurring from one census to another has been smooth and continuous.

The final issue in intercensal estimates involves the use of the forward/reverse survival rate method for producing detailed intercensal estimates. The forward/reverse survival rate method is usually applied to the two most recent censuses in order to derive a set of detailed estimates for the mid-census period. Forward survival rates are applied to the base population enumerated in the second to last census, and the results are then averaged with those obtained by the application of reverse survival rates applied to the most recent census. The forward/reverse survival rate method overcomes one of the disadvantages of simple interpolation strategies in the production of intercensal estimates in that detailed population estimates are made possible. The additional detail, however, comes at a cost. Data requirements for the forward/reverse survival rate method are more rigorous than those for simple interpolation methods, and the calculations are somewhat more difficult. Still, the forward/reverse survival rate method provides a relatively simple way to produce a detailed set of intercensal population estimates.

CHAPTER 5
ERRORS AND UNCERTAINTIES

## Scope of the Problem

It is critical to know whether the maps that can easily be created on GIS video screens are sound foundations for inference and decisions. As Michael Goodchild has noted often, statistics for spatial data is one of the GIS community's most pressing research needs. There are no confidence limits for the measures created with geographic information systems. Little is known about how error propagates through geographic information systems when maps are combined and permuted and when scales are changed. If a GIS is to be a useful tool, it must provide a measure of how much confidence can be placed in its products.

Physical scientists commonly use quality assurance measures to isolate biases in sampling, modeling and related analytical methods. Most common are those methods that yield quantitative, repeatable measures of the precision, accuracy, completeness, representativeness, and comparability of a data set. These measures are intended to provide adequate confidence that a data set satisfies user requirements. The same methodological constructs can be applied to digital geographic databases.

There is a lack of models of spatial data on which to build such things as indices of uncertainty, confidence limits, and statistical tests. Expecting a GIS that produces measures of uncertainty or confidence limits with all of its products is reasonable, so that in

measuring the area of a polygon, it would simultaneously produce the measure of uncertainty in that area. That is not available in any GIS, and yet that capability would be immensely valuable.

What will solve this problem is a series of statistical tools based on stochastic processes that can be used to model error. The analogy to conventional measures is that the GIS lacks the equivalent to the normal distribution used to put confidence limits on results and to construct things like t-tests. None of the equivalents is available in the GIS field presently.

Other concerns are aggregation and generalization, and variable reporting zones. It is amazingly difficult in population determination, because of changing reporting zones, and irregularly produced data to do any analysis that is spread out over a significant period, because during that period it is almost inevitable that reporting zones will have changed. So there has been a disturbing trend in recent years toward using zip code boundaries as reporting zones that have much poorer quality control than any of the zones that the census regularly uses.

Precision

Precision is the degree of mutual agreement among individual data sets made under prescribed conditions about position, factual entry or temporal validity. For example, two digital technicians, after an identical training program, can be instructed to capture the same set of point data with the same equipment and yield different results. The measured precision then is the variation between technicians. Precision can be quite useful in the identification of systematic error.

Accuracy

Accuracy is the measured degree to which the recorded value matches the true or referenced value. Bias is typically a measure of accuracy expressed as difference between a sample average and the true or referenced value. Acquisition of true or even reference values is sometimes difficult in a geographic database. Point data sets most easily lend themselves to an accuracy measure. Point data bias measures may aid in the detection of a systematic error with hardware or digitizing methods in either an x or y direction.

Accuracy, or the reduction of uncertainty, includes the following processes: data capture, processing, and compatibility. Each of these levels can introduce error and skew results. Three factors affect the usefulness of the processed data: how was the data captured, what kind of processing has been done, and how applicable is the processed data to the original purpose of the data. The goal for any system is to reduce uncertainty to such a level where decision-making becomes more effective.

Data capture and entry are two of the first steps in implementing a GIS. However, as data are entered into a system, three areas of uncertainty are encountered: locational, classification, and value. Locational uncertainty is directly related to place and scale is the deciding factor. As scales change, relative locations also change.

Uncertainty in attribute classification is another consideration for data capture. For instance, parcel ownership is unified in a spatial context but soils or vegetation for the same parcel are rarely homogeneous. In these cases classification then becomes an issue. In the example of vegetation, many sources of uncertainty can occur. Boundaries for physical attributes, as opposed to political attributes,

can be subjective assumptions of the field observer. Data gathering techniques and guidelines need to be explicitly laid out before field work is done to ensure the highest possible certainty in data. The quadrangle conforms to National Map Accuracy Standards: not more than 10% of points on the map are in error of more than 1/30 inch, measured at the publication scale. This standard, however, applies to "well-defined" points such as road intersections and benchmarks. Error in the trace of linear features, therefore, may be uncontrolled by the accuracy standard.

Data Processing

Uncertainties can enter the system and propagate when analytical processes and modeling are done. If not fully understood, the most basic analytical processing such as changes in scale and projection, filtering and other image enhancement techniques can introduce errors. Each time a subsequent process is done, the errors are multiplied, eventually producing totally spurious results. Processing produces new data that is an artifact of the original data, i.e., no longer empirical. The greatest value of a GIS, the capacity to process data to yield information in a spatial context, is also the biggest danger. For example, a general site location is digitized from a standard USGS 1:24,000 topographic sheet. Enlarging the image using a common zoom feature might lead someone not familiar with the function to assume that a 1:6,000 scale could be produced with the associated increased accuracy. This is not an uncommon conceptual error made by GIS technicians. The new image could be used to place samplers or dosimeters improperly due to accuracy problems. Comprehensive documentation is required of not only the database, but also of the analytical functions performed on the database. Without such

documentation, the uncertainties will be so high that decisions based on the resulting information might be incorrect.

Data Compatibility

Another factor affecting the accuracy of a GIS is the compatibility of the application to the original data. Additional applications are often found for existing databases (Antenucci, 1979; Tomlinson, 1982), but these databases are not necessarily suited to the new analyses. Data entered into a system should be collected in a fashion consistent with a GIS. However, this is often not so, and the quality of the data set may be lacking. A data set can contain problems with coverage, content, and age that are undocumented, but data entered into a system is usually assumed to be correct by the user. The user should avoid data that has not been documented in respect to date, sampling method, and definition of variables. The entire GIS analysis can be corrupted by a single inaccurate data layer. Clearly resolution and documentation are the main determinants of accuracy. The magnitude of uncertainty in the information derived from a GIS is inversely proportional to the spatial resolution and the attribute data documentation.

Federal, state, local and private organizations constantly produce volumes of spatial data. This creates an information management problem that must be solved by managing data about data in more effective ways than current practices. Metadata is data that describes other data. Some important types of metadata are data structure, meaning, and quality. Spatial data quality is very important to this study and includes these factors: positional accuracy, lineage, attribute accuracy, temporal accuracy, logical consistency, and completeness. Each element in a spatial database has the potential of being at a

different level of accuracy as more diverse sources are combined in common databases.

The GIS technology user should require levels of confidence or figures of merit to be issued with a digital database. Lacking such measures practitioners have expended considerable effort to evaluate the quality of third-party supplied digital data. Often publicly available data are discarded as a source because of a lack of documented quality indicators or because the data were developed for a non related application. Technologists continue to develop customized databases of limited value to potential peer users because of variation in resolution and data aggregations. The positional, temporal, and factual quality of a data set should be evaluated with a series of quantitative repeatable standard quality measures. These methods will be useful to detect bias and the presence of systematic and random errors that may have been introduced during the construction of the database.

Generalization

The information content of an attribute category is inversely proportional to the level of generalization. As attribute categories increase in generality (i.e., males versus females as opposed to Hispanic females versus African-American females), the less uncertainty there is but the information content is reduced. As a scale is increased there is less uncertainty regarding locational accuracy and information content. However, general data are of less economic and often scientific value than specific data. So it is a tradeoff based on economics and the detail necessary to answer a query. The accuracy required should be examined very carefully before setting up a GIS database. Decreasing the resolution of a data set is much easier than to increase it after acquisition.

Generalization is an inherent characteristic of all geographical data. For instance, as the scale of a map is decreased, there is less physical space to represent the geographic features of a region: as the process continues, such features are exaggerated in size to be distinguishable at a smaller scale. As geographic features "fight" for representation in the reduced map space, some features will need to be eliminated and those remaining may be further simplified, smoothed, displaced, aggregated or enhanced. In the extreme case, the map loses its geometric properties and becomes a caricature. Most of these generalization operations are carried out by trained cartographers, i.e., using manual methods. Some generalization operations, however, can also be carried out automatically, i.e., by using mathematical algorithms.

The generalization effects inherent in the data held in a GIS usually result either from the digitizing of previously generalized paper maps or from the subsequent use of generalization algorithms within the GIS (typically on data that have previously been generalized). From these observations, then, it can be inferred that all data within a GIS are generalized. Since the statistical and geometric properties of spatial data in a GIS are affected by the degree of generalization applied, so too are the results of any spatial manipulations performed on the data. Yet, though this is a fundamental area of research, very little quantitative work has been done in this field.

Generalization effects can manifest themselves in different ways, for example, by length and angularity change, by elimination of features or by displacement of features. Manual generalization effects embedded in paper maps are imported into a GIS when the paper maps are converted into a digital form.

Even when individual features show apparently small generalization effects, the cumulative impact on a GIS analysis using several different generalized features in combination can be much larger. This was demonstrated by Joao et al. (1992). The authors carried out an analysis to determine the length of road within 100 meters of a river at three different scales of source map representation. Between two of the scales the results differed by as much as 65%, caused by the simplification of the features and the changes in the relative position of the river and the road.

It is a reasonable inference that all users of GIS should be aware of the problems of generalization when carrying out an analysis and should seek to work with the largest scale maps available. Besides the generalization effects that are "locked into" paper maps, automated generalization algorithms can further generalize data within a GIS. The differences caused by generalizing automatically are less pronounced than those produced by manual generalization, but they may be in addition to generalization already carried out on and "locked into" the paper version of the map. Joao et al. (1992) presented evidence that generalization effects are typically greater in manually digitized paper maps than in those produced by at least one particular automated mathematical algorithm (the Douglas-Peucker algorithm--one of the most common data reduction or pseudo generalization algorithms available in GIS). This was because the Douglas-Peucker algorithm (Douglas and Peucker, 1973) only filters the high frequency components, causing a reduction in detail of the line without the displacement that manually generalized lines typically suffer.

Although automated generalization might appear to do a better job in statistical terms, graphically the result may be inferior because automated generalization only looks at individual lines "one at a time";

manual generalization typically considers more of the map as a whole, the cartographer moving one feature to give space to another. Automated generalization methods are poor in the maintenance of relationships between features unless these are explicitly specified through a topological statement, typically polygon adjacency. For example, the Douglas-Peucker algorithm cannot of itself cope with a situation when a contour and a river to be generalized are still in the same position relative to each other after generalization (i.e., the river crossing the contour at exactly the same point). The solution is to store the intersection point between the two features. Rhind's rule (Rhind, 1973), therefore, is to do overlay of all related features before generalization, to maintain fixed points in automated generalization. In general terms, where care is taken in maintaining the geographic relationship between features, GIS analysis is liable to be more reliable, if less graphically pleasing, when using lines that resulted from automated generalization. However, if any pictorial analysis is needed, then maps derived from manual generalization would often be better because they would be clearer visually.

It is manifestly important that automated generalization algorithms do a "good job," both statistically and graphically. This has been reflected in the recent development of new algorithms that try to mimic the results produced by manual generalization methods. For example, Li and Openshaw (1992) proposed a new set of algorithms for line simplification that is better than existing counterparts at producing results equivalent to those produced by manual methods. The authors called this approach "the natural principle of objective generalization." Other researchers (Muller and Zeshen, 1992) have also proposed new "natural" algorithms for the generalization of area-patches. The development of these new types of algorithms reflects a

dissatisfaction with existing pseudo generalization algorithms within the GIS community. Visvalingam and Whyatt (1990), for example, pointed out other problems arising from the widely used Douglas-Peucker algorithm, such as the creation of lines that are too "spiky."

The generalization problem is widely acknowledged in academic studies. Other researchers, such as Goodchild (1980), Blakemore (1983) and Muller (1991), have assessed qualitatively the impact of generalization effects on data quality. Goodchild (1980) described some types of data errors that can occur in both raster and vector systems due to the effects of generalization, such as in area and point estimation, and in the creation of many spurious polygons. Blakemore (1983) considered that generalization effects create potentially serious geocoding and retrieval errors. Muller (1991) pointed out that generalization can alter data quality by decreasing location and attribute accuracy, and by affecting completeness and consistency within the data. Yet, despite the acknowledgment of the generalization problem within academic circles, no mechanisms of any kind have been developed or proposed within commercial GIS to reduce the problems arising from generalization or even to warn users of its likely effects.

While the spread of GIS may increase the use of spatial data, it can also result in many relatively inexperienced users trying to produce adequate results without fully understanding the implications of generalization. Quality issues relating to precision, accuracy, completeness, representativeness, and comparability of the data all are affected by the degree of generalization. Since a map is a reduced representation of the surface features, the scale of the map largely determines the amount of information that can be stored either graphically or as attributes. The consequence of scale reduction is generalization; the smaller the scale the greater the degree of

generalization. Information contained in the map has two components, location and meaning. Generalization affects both. In the context of the digital geographic databases, cartographic generalization is the collective modifications made during map preparation and automation. Methods of generalization include selection, simplification (elimination), combination, exaggeration and displacement, and classification. The results of each of these methods can affect an end user's satisfaction with a given database. For features that possess areal extent such as soils and land use, a minimum mapping unit is typically applied.

Value Uncertainty

Value uncertainty is based upon absolute spatial relationships. For instance, linear features on a map such as jeep trails, highways, interstate highways, creeks, rivers, railroads, and others are represented by lines of equal width. This can be conceptually misleading when their respective dimensions have importance. Although their relative location is correct, their absolute spatial value is uncertain due to cartographic modeling constraints. Resolution is the critical factor: the higher the resolution, the more accurate the value will be in terms of both relative and absolute location.

## Dealing with Error

GIS is, above all else, an integrating technology that allows, encourages and expects users to bring together data from many different sources through the unifying medium of geography. It is widely known that geographical information is subject to error and uncertainty. This applies to both the map base data, due to data capture error and generalization effects, and the associated attribute information, due to

a mix of measurement error, sampling error, and the effects of categorization and aggregation. Uncertainty and error are, therefore, endemic characteristics of all spatial information, to which additional contributions are made by applying various GIS models to represent the world. Additionally, many GIS operators add their own errors to data that already contain error. This is only a problem if users ignore the effects of handling uncertain spatial information when they make use of it. The problem is that virtually all the commercial GIS in the world offer no error handling facilities and nearly all the world's supply of spatial data has been stored without any detailed error tagging. The word "error" is used here in the statistical sense of uncertainty as to what the precise value of a coordinate really is. In many spatial databases the values for the millions of coordinates may be precise to 5, 6, 7 or even 10 digits, or they may not be quite so precise. It may not matter much or it may be catastrophic. Once it is realized that this uncertainty characterizes all aspects of spatial information, then there is cause for concern, especially as the prevailing wisdom is to ignore it. There are various causes of uncertainty, most of which cannot be simply removed by improved data capture. There is also the added problem that the objects being represented in GIS are themselves difficult to define precisely, and possess what Burrough (1986) refers to as "unseen errors" and "natural spatial variation" (p. 132). In practice, these natural sources of variation are augmented by probably not inconsiderable amounts of artificially generated or unnatural uncertainties added to the data by the GIS process itself. Major sources of error in GIS are: data capture error, ageing of data, representational, generalization and simplification processes, and GIS operations. One interesting fact about errors in spatial information is that often they may not matter. Under such circumstances, it is

convenient to ignore them by pretending that they do not exist. The problem is that whether errors matter is dependent on both context and application. It is neither safe to assume a priori, on the basis of little or no relevant knowledge, that errors do not matter, nor that they do matter to the extent that an application or a spatial decision support task might be put in jeopardy. In most cases the user does not know and, worse still from a scientific point of view, has no obvious way of finding out. It follows, therefore, that the absence of error handling methods in existing GIS is a grave functional deficiency. Some consequences of ignoring errors in GIS are: selection of "wrong" locations, identification of false patterns, derivation of erroneous data, failure to find relationships that matter, and poor decisions that cost money. All data contain errors and it is important to discover how to cope with uncertain spatial information rather than to either pretend it does not exist or revert to non-GIS tools that contain even larger amounts or error (Openshaw, 1989).

Solving the Problem

The National Center for Geographic Information and Analysis (NCGIA) recognized the importance of error in spatial databases and based their first initiative on this theme. This initiative is now "closed", but it seems that most of the problems that were identified have not been resolved. This initiative mainly focused on raising the level of user awareness of the problem; (Goodchild and Gopal, 1989). Research attention was then focused on error in remotely-sensed data rather than the more general GIS case (Goodchild, Guoging and Shiren, 1992), and virtually no attention was given to the problems of error simulation and propagation. The initial approach to studying error propagation in GIS was based on Openshaw (1989), who identified a

general purpose Monte-Carlo algorithm. The idea is deceptively simple and generally applicable. It involves "wobbling" point/line segments to reflect best estimates of the underlying uncertainties in their positions in space. This process is performed upon each data source in turn, then some arbitrary sequence of GIS operations is applied. Eventually a result, or output coverage, is obtained. The "wobbling" process would be repeated M times; for example, M might be 99. The variation in the results would be summarized as one- or two-dimensional distributions and used to provide approximate "confidence regions". In a polygon overlay application the results of each data wobble could be rastered, with counts kept at the pixel level of the frequency of each pixel being an acceptable site. It is fairly easy to determine "safe" regions on the map. Openshaw et al. (1991) provide an illustration of this approach in practice.

The problems with this Monte-Carlo method include these items: (a) it is computer-intensive, taking orders of magnitude more time, since the GIS operations may have to be repeated many times; (b) the "wobbling" of the line data can cause topological problems in that line segments may cross and the topology may be damaged or changed; and (c) the error model for the wobbling is not defined and needs to be identified. In some ways these problems are now less relevant. The additional computation times are largely irrelevant as a result of trends towards faster hardware and multiple CPU workstations, and also because of the prospect of reduced numbers of Monte-Carlo re-runs due to the use of sequential significance tests. The topology problem can be overcome by switching from vector to raster for the wobbling and then back to vector, or else by remaining in the raster world. If the former strategy is used, Brunsdon and Carver (1993) and Carver and Brunsdon (1993) provide the basis for an optimal conversion. The lack of an

error model is a more difficult problem. There are two possible strategies: either regard the wobbling as a form of sensitivity analysis in which the impact of relatively small data perturbations is to be investigated without assuming they reflect levels of real error; or develop explicit error models that seek to represent plausible levels of measurement uncertainty. It is argued that this Monte-Carlo approach to error simulation as a form of sensitivity analysis is now a practical proposition. Finally, it should be noted that, as Openshaw et al. (1991) show, errors in a GIS environment certainly propagate and do not necessarily cancel each other out. They cannot safely be ignored. Indeed, the only "safe" way to use GIS is to ascertain estimates of the likely effects of errors on a specific application before deciding whether they matter. Furthermore, error propagation in GIS goes beyond map data. There can easily be interactions between uncertainties in the positional information and various attribute errors. Also, the use of spatial models in decision support introduces the additional need to include model generated uncertainties (Openshaw, 1979) as well as digital map related ones. The Monte-Carlo approach can, in principle, cope with both.

It soon became apparent that if the error simulation problem involved only overlay, then a much simpler approach was possible. The concept of error epsilons around lines is an old idea; see for example Perkal (1966). The method creates buffers around a line with a width designed to reflect positional uncertainty in the location of the line due to measurement and other errors. Brunsdon et al. (1990) have generalized and extended this procedure to allow (a) the buffer width to continuously vary, and (b) quantification of the uncertainty buffer as a piecewise quantic function. Overlay operations can then be specified that combine, using probability theory, the buffer regions around

different line coverages. This has been implemented as a set of commands for ARC/INFO (see Carver and Brunsdon, 1992; Carver and Openshaw, 1990). These probabilistic epsilon error commands provide a useful toolkit for handling overlay problems. It is sophisticated in that the uncertainty region can change along a line segment. The uncertainty information can be stored as a data attribute. However, there still remains the problem of how to generate realistic estimates of positional uncertainty when none was stored with the spatial data when it was captured. The problem of error simulation is now the biggest unresolved task in GIS. By comparison, it now seems that developing error handlers is relatively straightforward, albeit computer intensive.

The focus of attention is now on developing a means of predicting the likely magnitude of measurement error in line data that was captured without any such information. It is assumed here that point data error prediction is of little interest because such errors are probably well represented by a normal distribution assumption using variance estimates that are a function of map scale. It should be fairly easy to determine these parameters via experimental methods.

Measurement error in this context is that associated with digitizing cartographic line features on paper maps. This is regarded as the most general problem that is worthy of attention. It is argued that any model that can cope with this aspect should be able to handle other variants, such as measurement error in automatic digitizers and even raster-vector conversion. It is taken for granted that any such errors will be scale-dependent. The main problem is that in general we do not know which error model to apply, as this is rarely provided or indeed known in any detail. Another problem is that measurement error cannot be assumed to be either uniform over a whole map or even uniform

along a single line. It is obvious that some parts of a line are more difficult to digitize than others. Unfortunately, there is unlikely to be any simple relationship between the geometric complexity of a line and its error. For instance, it is possible that operators take more care and work more slowly with highly irregular line segments than they do with simple ones. On the other hand, cartographic generalization may be far more severe on the more complex line segments while others might be completely untouched.

One possibility for the measurement of digitizing error: the measured error is based on the perpendicular discrepancy of the digitized line with the true line at the node points on the true line. There may be some shortcomings, particularly when there are very long segments of digitized line, but these could be overcome by introducing pseudo-segments by subdividing very long segments. This has the advantage of being relatively economical to calculate, while still giving a reasonable measure of digitizing accuracy at regions along the cartographic line features. This can be further enhanced by also considering the errors of the digitized points in addition to those of the original line. A final specification must now be made, relating to the degree of error margin that should actually be allowed for when processing the map data. One possible choice might be in terms of confidence bands about the map line. With error information obtained on particular features within a line, it will be possible to find values that just exceed, say, 95% of all measured errors. These are effectively 95% confidence limits, determined empirically from measured error on map features of known complexity.

Completeness

Completeness is the degree to which all intended entries into a database have been encoded and/or digitized. Omissions of data do occur and in some cases erroneous entries can be made. Measures of completeness can be simply expressed as a percentage complete by data element category.

Representativeness

Representativeness is many times an elusive measure even for the most experienced statistician. Representativeness is the degree to which the recorded observation represents the study area or analysis area as a whole. Many times sample plots are used to characterize a large area without a measure of representativeness. Selected measures of representativeness should be considered only after considerable forethought due to cost and logistic concerns. Soil and vegetation geographic databases are examples of data sets which may require a high cost to measure representativeness of detailed sample plots.

Comparability

Comparability is the degree to which a database can be merged with others covering the same geographic area or containing similar thematic data sets of adjacent areas. This is especially a concern for users of multitemporal and multiscale data which has been prepared by different entities.

## Error in Digital Representations of Lines

Although geographic information systems store data in various ways, many define objects using a vector approach (Peuquet, 1984; White, 1984). When data are vector coded, objects such as road networks are

often defined by chains of points with associated topological linkages. Network distances are computed in vector databases by summing scaled interpoint distances within each chain.

Deucker (1975) has outlined a general theory of cartographic lines. The theory is especially pertinent to the treatment of line complexity. Chrisman (1982) has developed a preliminary theory for assessing cartographic error and its measurement. Other researchers (e.g. Muller, 1987) have addressed specific dimensions of error and have identified a series of error sources. These sources can be placed into the following categories:

Sampling Density

The Steinhaus paradox (Perkal, 1966; Buttenfield, 1985) shows that as smaller sampling intervals are used to measure a sinuous feature, it apparently increases in length. Sinuosity which was missed in previous intervals is now detectable, and is added to the previous length estimate. During digitizing, decisions are made about data encoding precision. These decisions affect distance calculations.

Source Material Scale

The scale at which a map is compiled can influence length measurement because generalization, in the form of decreased line complexity, increases as map scales become smaller (Gardiner, 1982; Maling, 1968).

Source Material Error

Dueker (1975) has described error types that occur on source maps. If the source contains error, its digital product will, at best, replicate those errors with a high degree of fidelity. It is likely, however, that additional error will be introduced during data capture.

Error in Data Capture

When documents are captured in digital vector form, encoding errors occur (Dueker, 1975). Jenks (1981) has identified errors that occur during manual encoding of irregular linear features.

Error in Data Storage and Processing

Chrisman (1984) argues for using good judgment in selecting the resolution for coordinate storage, and for the use of absolute, rather than incremental coordinates. Chrisman also describes other issues that must be weighed (e.g. real vs. integer, and decimal precision).

Topological Error

Ginsburgh and Hansen (1974) describe three common error types in studies using road networks. Their third type deals with topological problems arising from incomplete specification of internodal linkages. In their framework the topology of a network is verified by reciprocal enumeration.

Aggregation Error

Hillsman and Rhoda (1978), Goodchild (1979), and Casillas (1983) have shown that aggregating demand (e.g. households) to a single proximal point misrepresents the actual underlying distribution of demand and introduces error into analyses.

Also note that in some instances, other factors contribute to measurement error (e.g. document instability, projection). Each of the above factors can act alone, but error is often attributable to several factors acting in concert.

## Polygon Overlay Processing

The polygon overlay processing (POP) problem in GIS can be broken into two fundamental components. One is the determination of the geometric intersection of two geographic entities. The other is the assignment of attributes to the resulting composite entities. What makes the POP problem so attractive for the application of expert systems is the amount of domain-specific processing that often must be done. The fundamental operations of POP are rather simple, but as much as 80% of the code for some overlay systems is devoted to handling special cases, input, output, and book-keeping. Thus, there is a fair amount of knowledge pertaining to the geometric aspects of POP.

Modeling the management of multi-layer polygonal uncertainty affects a wide variety of data analysis operations in GIS. Among those operations affected are point-in-polygon, transacts, profiles, and calculation of distance, area, and direction. There are implications affecting the responses to spatial queries.

The spurious polygon problem continues to seriously affect the utility of GIS that rely on polygon overlay processing for integrating data layers. That problem can be encountered by using the wrong technique to construct radial lines in the sector-segment grid; specifically, use of the diagonal line drawing command. That command leaves spurious polygons around the grid origin. The user must interactively repair the grid, or properly construct a new one.

## Error Estimation for the Sector-Segment Grid

A standard 50-mile sector-segment grid used for offsite dose assessment consists of fourteen annular segments (0-1, 1-2, 2-3, 3-4, 4-

5, 5-6, 6-7, 7-8, 8-9, 9-10, 10-20, 20-30, 30-40, and 40-50 miles) and sixteen radial sectors centered on the major compass points for a total of 224 sector-segments. A typical carefully constructed grid was found to have an arithmetic mean of 0.13 percent error per sector-segment. The population of sector-segment errors is normally distributed with a standard deviation of 0.037. This means that a researcher interested in a sector-segment in the 40-50 mile annulus may include or exclude about 0.23 square miles of geography on average, and 68 percent of the sector-segments will range from 0.16 to 0.3 square miles in error. This may be significant depending upon the spatial distribution and density of the local population. The actual population residing within a sector-segment at the time of a decennial census is estimated by constructing the grid over the census boundaries and querying the attribute database for the total population within each sector-segment. To accomplish this, the GIS software assigns a portion of a partially-included census region to the corresponding sector-segment based on its area. The percentage of its area included in the sector-segment is applied to its population as a weighting factor. By repeating this process for all affected census regions and summing results, a population total for the sector-segment is generated. The problem is that people are seldom uniformly spatially distributed. This is another source of error. The estimated population will have errors associated with: 1) the census count itself, 2) the encoded census boundaries, 3) sector-segment grid construction, and 4) area-weighting of the population. Of these, categories 1 and 2 are beyond a user's control. The census count error is used in error propagation as reported by the Bureau of the Census for the decennial census of interest. The encoded census boundary error is assumed to be negligible. Category 3 is one that the user can attempt to deal with, and error due to category 4 can be minimized by use of the

smallest possible census boundaries, e.g., the block level for the 1990 census, in the analysis. Efforts at writing a code to run within ATLAS GIS to constuct a sector-segment grid based on geometry rather than user manipulation have not been successful. Such a code would streamline the analytical process, and eliminate the human error component. This is the most readily corrected error component. Inherent errors in line-drawing algorithms would not be affected. ATLAS GIS code is written using a script language similar to BASIC and dBASE programming tools. This allows the programmer to construct highly complex scripts that provide direct access to the core-level functionality of the GIS software. The problem is that geometric operators are not allowed in the program. Therefore attempts to construct a buffered grid based on geometrical concepts is impossible. This is possibly an exclusion to preclude anomalous results that may occur due to the on-the-fly latitude/longitude pseudo projection, or translation to true projections that would not support such a construct.

An estimate of the population error attributed to area-weighting can be attained by selecting all census boundaries intersected by the boundary lines of the sector-segment of interest (see Figure 5-1 and Figure 5-2). These are the area-weighted regions, and their population assignments can theoretically approach 100 percent error. A conservative treatment is to assign 100 percent error to each, sum for all intersected regions, and divide by the sector-segment total population for the fractional error. A more reasonable approach would be to assign 50 percent error in place of the 100 percent above and proceed as stated. Most GIS platforms use the same underlying structure. Consequently, errors and uncertainties described in this work are applicable to most systems (e.g., ATLAS GIS, ARC/INFO, etc.).

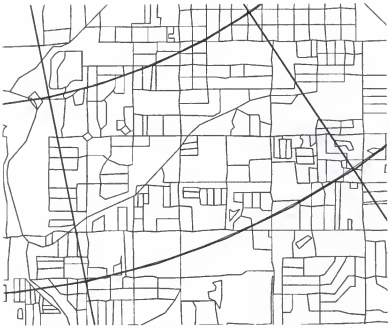

Figure 5-1 Inner Grid Sector-Segment (2 - 3 Miles) Over Census Blocks

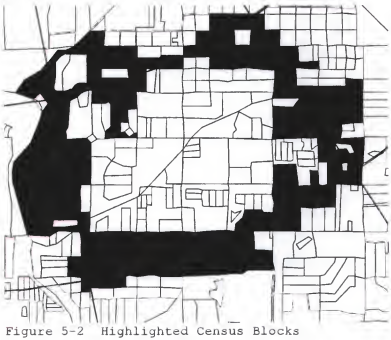

Figure 5-2 Highlighted Census Blocks Intersecting Grid Lines

An assessment of this description was applied to the sector-segment grid and census block populations used for the Pinellas Plant. Of the total population in the 50-mile radius of 2,529,869 persons for 1990, only 20,998 reside within blocks intersecting grid lines. This is 0.83 percent for the entire grid. Naturally this is a slight underestimate of the error for small sector-segments near the grid origin and progresses to a conservative estimate for outer sector-segments whose affected to total block ratio is smaller. Also, note that errors would be higher for larger census boundaries, i.e., block groups and tracts, but those levels are not recommended for population analyses by this author.

Martin Marietta staff calculated an average individual dose of 7.5 E-7 mSv/yr in the 50-mile radius based on CAP88-PC analysis of tritium and Kr-85 emissions. They applied that average individual dose to the entire grid population (supplied by this research) of 2,529,869 persons for a collective dose of 1.897 E-3 person-Sv (recall the dose equations of Chapter 1). Considering grid construction and area-weighting allocation errors in a tandem gives an indication of their contribution to the total dose error. The maximum grid error is 0.0013 + 0.0037 (one sigma). The maximum error in over projecting the grid is represented by 1 + 0.001337 and the maximum error in under projecting (minimizing) the grid is 1 - 0.001337. The population error is reasonably estimated as plus or minus half the total population of the intersected blocks, or +/- (20,998/2). This component yields a population range of 2,519,370 to 2,540,368. The maximum error bounds are 2,540,368 multiplied by 1.001337, to 2,519,370 multiplied by 0.998663. The result is 2,543,764 to 2,516,002 or a range of 27,762. Application of these bounds to the dose calculation yields a range of 1.908 to 1.887 person-Sv. The small error range is due to the fact that this particular sector-segment grid

is centered on an area of high population density; a little more than half of the entire grid is over water crating no impact on the population error; and the outer 50-mile ring, which bears the full impact when the inner sector-segments are summed, lays over predominantly low-density areas. The population error must increase for any analogous analysis where population densities are more uniform and/or larger census entities (block groups or tracts) are used.

CHAPTER 6
APPLICATION OF NEW TECHNIQUES TO ENVIRONMENTAL PROBLEMS

## Overview

Environmental impact studies require spatial analysis or manipulation of population distribution, ecological factors, and particular sites of importance. GIS can offer a more objective means of weighting the relevant factors in the site-selection phase for a radioactive waste disposal facility, or any other environmentally and socially sensitive disposal site. GIS confers the ability to rapidly develop alternative scenarios by reassigning numeric significance to the different criteria, such as groundwater levels, average annual precipitation, geological constraints to facility construction, population density boundaries, and more. The quality of GIS that makes this technology so powerful is how it enables the user to combine independent data sets from a variety of separate disciplines, and thereby solve complex problems that can be difficult to resolve by traditional means. Commonly, the GIS is configured to handle data in two dimensions as static elements. Environmental sciences frequently make extensive use of the third vertical space dimension. Also, many problems also involve the dimension of time, for example the modeling of radioactive plumes. The development of higher-dimensional data structures capable of storing very large data sets is required. Methods of interpolation in up to four dimensions are also needed to augment incomplete data sets. Current GIS structures are limited to simple

empirical modeling. However, many environmental models are complex deterministic or stochastic models. The solution of sets of partial differential equations in a distributed model using finite element techniques is not possible within a current GIS. As a consequence of integrating spatial data containing inherent errors, techniques for the analysis of error propagation are also required (see previous chapter).

Geographic Criteria for the Siting of Low-Level Waste Disposal Sites

Geographic siting criteria for near-surface disposal of low-level radioactive waste (LLW) include population distribution, land use, natural resources, geomorphology, and other spatially oriented factors. The geographic and other technical site selection requirements for near-surface LLW disposal are presented in paragraph 61.50(a) of 10 CFR Part 61. These requirements specifically deal with conditions that could affect long-term site stability and waste isolation. The site suitability requirements may eliminate from consideration land that has certain unfavorable hydrologic, geologic, land-use, and demographic conditions that could adversely affect the site and its surrounding environment. This discussion addresses the demographic requirement and its evaluation.

The prospective site should be located in an area of low population density where the potential for future population growth is estimated to be quite limited. The determination of population densities with a GIS is discussed in this study, but definitions of "low" can be arbitrary. There is an additional stipulation to assess future population growth, i.e., a necessity for population projections. That projection methodology is presented in this document. Also, this site should be at least 2 kilometers from the residential property

limits of the nearest existing urban community. Once site boundaries and existing urban community boundaries are input to the GIS, it is a simple task to find conflicts within the 2-kilometer zone.

Applicable State and local land use plans and regulations (including zoning ordinances) should be fully scrutinized to verify that there are no conflicting regulations or development plans near the site. Existing residentially zoned or planned land uses are considered conflicting and prohibitive factors within a 2-kilometer radius of the candidate site. In addition, local and State authorities should be consulted for information on planned highway construction in the vicinity of the site to avoid interference with the operation of the site, and to decide whether there will be adequate access to the site in terms of future highways and land use.

## Environmental Modeling

Environmental process simulations would benefit from the development of GIS capable of coping with the additional dimensions of vertical space and time and having extended spatial modeling facilities. For example, consider a distributed, physically-based process model predicting the state of the atmosphere in three-dimensional space and time. Existing GIS is of little use other than peripherally in this modeling process because of limitations in the GIS itself. The limitations of current GIS for environmental modeling become apparent when one considers the types of models involved and their associated data. Various environmental models are discussed in depth by Goodchild, Parks and Steyaert (1993) and Farmer and Rycroft (1991). A problem in oceanographic modeling is the synthesis of time-dependent ship and satellite data in order to improve numerical models of the ocean. The

World Ocean Circulation Experiment (WOCE) is studying the present state of the ocean circulation through observations and the development of models (Smythe-Wright, 1991) to provide a basis for predicting climate change. The traditional method of obtaining a spatial physical description of the ocean is to conduct a hydrographic survey, in which a ship carries out vertical profile measurements at a number of discrete locations (typically separated by about 20 km). Three-dimensional fields of measured variables are computed. The use of satellite sensors such as the Advanced Very High Resolution Radiometer (AVHRR) enables the synoptic mapping of variables at a single level (the surface). There is a need to investigate the relationship between remotely-sensed surface signatures and subsurface ocean features to combine these data sets.

Such detailed environmental studies demonstrate that, in contrast to the two-dimensional static data sets handled by most current GIS, it is necessary to link models to data sets of up to four dimensions. These data sets are generated by many different types of sensors, and include remotely-sensed data that provide a two-dimensional synoptic view. These data inconjunction with smoothing algorithms permit extrapolation from site-specific measurements to wider regions. Data from different sensors may be of different dimensionalities, resolutions and accuracies, but they must be integrated. The size of data sets compounds from two to three to four dimesions of data such that model output files are gigabytes in size. Data may also be sparse in some dimensions compared with others, e.g., frequent sampling in the vertical and sparse sampling in the horizontal.

Many environmental models are complex deterministic or stochastic process models, which may operate in three-dimensional space and time. Deterministic process models produce output parameters by solving differential equations describing knwon chemical or physical processes.

The models may be distributed, finite element or finite difference models. The spatial context in which the process operates is subdivided into regions of relatively homogeneous modeled parameters, but heterogenous between adjacent regions. The output from one region forms the input to adjacent regions, and vice versa. Stochastic process models simulate the average outcome of many events rather than the outcome of just one event. Linked submodels, such as a hydrodynamic model describing water movement used to drive an ecological model describing water quality, combine to form a more powerful composite model.

Spatiotemporal GIS Techniques

Current GIS generally based on digital maps handle data of two spatial dimensions that are static in time. They typically permit only simple empirical modeling, and the complex environmental models are far beyond their scope. The solution is a GIS "shell" capable of supporting higher dimensional data sets and having extended facilities for mathematical modeling.

Higher-dimensional data are often sparse in one or more dimensions, so there is a need for the GIS to incorporate methods of interpolation in up to four dimensions to allow the user to interpolate between sample points.

Volume and object data

One method is to treat the spatial and time axes in the same way in the volume data model. The volume data are assumed to be embedded in four-dimensional Euclidean space. For example, a four-dimensional data set can be displayed as a time series of three-dimensional (x,y,z) volumes, or alternatively as a series of (x,y) and time plots at fixed value of z. Different ways of treating space and time will sometimes

have to co-exist, e.g., the use of different interpolation algorithms for space and time.

Hierarchies often allow data compression of volume such as undefined volumes within sparsely-filled four-dimensional data sets. Hierarchies also provide a simple way of producing scale-variable volume data. Often four-dimensional data sets are very large and global views are attainable only at low resolution, but there is usually a requirement to view some part of the data set at high resolution (Gahegan, 1989, Callen et al., 1986).

Objects in environmental science data sets, such as radioactive plumes or meterological fronts, can vary in up to four dimensions. Scientists need to be able to define, store, manipulate and display objects and their attributes, and relate them to the volume database (see, for example, Worboys, 1992).

Interpolation in sparse volume sets

A feature of data in higher-dimensional spaces is that they are often sparse in some dimensions compared with others. Geological applications involving the use of borehole and well data provide well-known examples, as sampling is typically dense along each borehole's vertical profile, but the boreholes are distributed relatively sparsely in the horizontal plane. Another example is provided by the oceanography data set. In order to prevent the user drawing inaccurate inferences from sparse data, it is an advantage if the interpolation procedure is capable of providing error estimates on the interpolated values. An important method of interpolation falling into this category is kriging. This is a form of weighted local averaging based on the theory of regionalized variables (Olea, 1975). It is an optimal interpolator in the sense that each interpolated value is unbiased and its estimated variance is a minimum.

## Census of Agriculture Data

The census of agriculture is required by law under Title 13, United States Code, sections 142(a) and 191 which stipulates that a census be taken every fifth year after 1983 covering the prior year. The data are comprehensive and published on CD-ROM for efficient input into systems such as GIS. An example of how such data could be used follows.

In NUREG/CR-2850 PNL-4221 Vol. 10, dose commitments were tabulated for the total population and the 4 population subgroups displayed in Table 6-1.

Table 6-1 Standard Population Subgroups

| CLASSIFICATION | AGE (YEARS) | PERCENTAGE APPLIED |
|---|---|---|
| infant | 0-1 | 1.44 |
| child | 1-11 | 16 |
| teenager | 11-17 | 11.7 |
| adult | $\geq$ 17 | 70.9 |
| Total | | 100.04 |

These tabulations apply to those living between 2 and 80 km from each of 72 reactor sites. The age groupings were derived from population totals by applying the aforementioned factors to all site populations in the U.S. The factors were obtained from reference USBC, 1975. The population distributions were supplied by the NRC as an updated reduction of the 1980 census (Sinsgalli, 1992). Major metropolitan areas and Standard Metropolitan Statistical Areas (SMSA) extrapolated to 1988 populations were also listed. This document series could be improved using techniques described herein to provide new input data as listed.

1. population distributions for each site using either an English or metric grid based on the 1990 census using GIS technology to manipulate TIGER and P.L. 94-171 data (this scrutinizes the population down to the census block level which number in the thousands per typical county);

2. population projections using Bureau of the Census data or statistical abstracts generated by each state's own agencies which may be more accurate; and

3. age breakdowns as recorded by the Census Bureau and assigned to sector-segments. These are coded in the database as <5 yrs, 5-17, 18-20, 21-24, 25-44, 45-54, etc. The age breakdowns used in the 1988 publication may be "set in stone" for proper input into mathematical models contained in computer programs based on Regulatory Guide 1.109. If that is the case, percentages applied to these finer age classes for each sector-segment will still yield much more accurate figures than previously estimated. Breakdowns by single years are available in Summary Tape File 3A.

Dose commitments for each age and organ were tabulated for the water pathway and the airborne pathway. Input data for average production rates of vegetable crops and animal products within 80 km radii were based on the respective statewide average. Access to 1992 Census of Agriculture data on CD ROM is available by zip code. These data can be provided by sector-segment since the computer program automatically applies weighted fractions of zip code data to the sector-segments.

The 1988 report used the atmospheric dispersion modeling program XOQDOQ which required input of population data from a grid of 160 sector-segments with sectors centered on the 16 major compass points and segment rings of 2-3, 3-4, 4-6, 6-9, 9-14, 14-20, 20-30, 30-40, 40-60, and 60-80 km. The GIS techniques described herein are well-suited to supply the input data.

Census of Agriculture data can be used to generate site-specific data that may be used in the semi-annual dose assessment reports to the Nuclear Regulatory Commission (NRC). These reports must specify the principal radionuclides and the activities of each as released to unrestricted areas in liquid and airborne effluents. The facility must

provide sufficient information to estimate annual radiation doses to the public resulting from effluent releases. Semi-annual reports utilize the methodology of U.S. Nuclear Regulatory Guide 1.109. That document is intended as a framework for helping health physicists establish off-site radiation doses to the general public. The regulatory guide presents pathway equations and default parameters conservatively applicable to the entire continental United States. The NRC encourages use of site-specific data to replace the default values from Regulatory Guide 1.109, and generate more accurate estimates for the region. Local agricultural data indicates the predominent crops harvested in the area and their yields. When the population and typical consumption rates are factored in, one can determine whether a particular agricultural product is consumed at the maximum rate, or is diluted by non-contaminated products from outside areas.

## Case Study: The Cotter Mill Problem

The case study that follows illustrates the utility of GIS for spatially locating data points and comparing them to other data sets. In this case, the data points are mainly soil samples (with depth data) taken on the properties of plaintiffs in an ongoing litigation. The comparison data sets are post facto background values derived from literature, and data from a mineral survey. GIS technology was used to locate data by latitude/longitude coordinates and address matching. The data shown in the enclosed example map (Figure 6-1) and similar maps for other elements could also be viewed concurrently on the computer screen using the GIS layering feature.

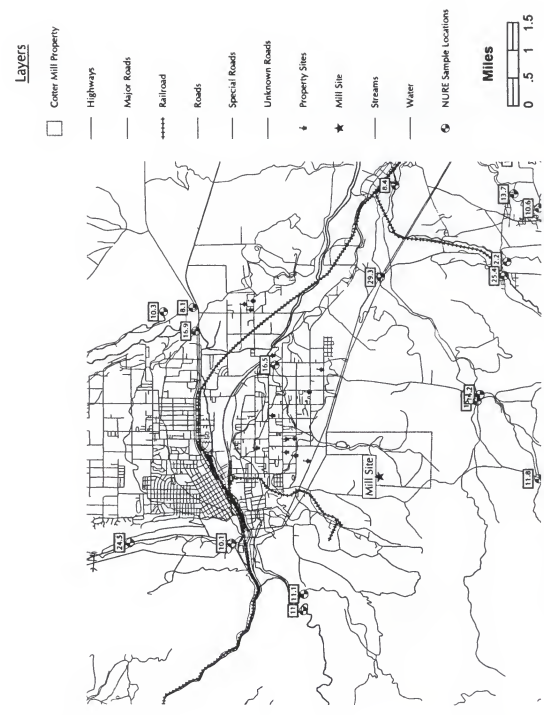


Figure 6-1 NURE Results for Thorium Near Cotter Mill

Mining and milling operations continually add heavy metal and metalloid contaminants to the area surrounding Cotter Mill near Lincoln

Park, Colorado. This often includes those considered undesirable, or toxic if concentrated, e.g., arsenic, selenium, and molybdenum. Contaminants of concern in this effort are arsenic, cobalt, copper, molybdenum, nickel, selenium, radium, and uranium. Chemical properties dictate the atmospheric transport of metals. Of interest in this study are some volatile "metalloids"; specifically, selenium and arsenic. Those elements can be transported in a gaseous form or enriched on particles, whereas other metals such as cadmium, lead, and zinc are transported only in the particle phase and may travel long distances before deposition to land (Haygarth and Jones, 1992). Some metals may adhere to crop surfaces, and may be absorbed via foliage as gaseous forms.

> There is growing evidence to suggest that microbially mediated processes may be impeded in soils at heavy metal concentrations somewhat above background. This may be of regional importance where atmospheric deposition has elevated the heavy metal burden of surface soils. (p. 137)

Uranium and radium are reported in the analyses in units of parts per million (ppm) as well as activity concentration, that is, nuclear transitions per unit time per unit mass (such as pCi/g). Uranium series radionuclides are assessed in terms of gross alpha and beta emissions.

Summary of Radioactivity and Contaminant Backgrounds

Mining and milling activities cause alterations in the geochemical balances that had become established through time.

> In semiarid or arid regions, such as prevail in most of the Western energy regions, the development of a balanced geochemical system requires a very long period of time. Because the amount of water from precipitation generally is small, rocks and minerals in the soil weather slowly, and organic deposition from the sparse vegetation occurs at a slow rate. These factors contribute to the development of a fragile ecosystem that is susceptible to rapid degradation by the disturbances of industrial development. (Ebens and Shacklette, 1982, p. 23)

That same report includes a study in which materials transported to the surface through mining activities had a pronounced effect on soil chemistry. Specifically, oil shale was found to be richer in arsenic, selenium, and molybdenum than were the soils from the same area in the Piceance Creek Basin of Colorado. Background values in ppm used in this report are displayed in Table 6-2.

Table 6-2 Trace Element Background Levels in Soil

| ELEMENT | SOIL CONCENTRATION (PPM) |
|---|---|
| arsenic | 7.2 |
| cobalt | 9.1 |
| copper | 25 |
| molybdenum | 0.97 |
| nickel | 19 |
| radium-226 | 6.6E-7 |
| selenium | 0.39 |
| uranium-238 | 1.8 |

These values were selected from 14 publications (Ebens and Shacklette, 1982; Shacklette and Boerngen, 1984; Kabata-Pendias and Pendias, 1984; Adriano, 1986; Bowen, 1979; Bowen, 1966; Rose, 1979; Aubert and Pinta, 1977; Swaine, 1955; EPA, 1982; NCRP, 1987; Lowder, 1964; Vinogradov, 1959; and Bear, 1964). None of the background values used in this report is the lowest found in the literature (lower values would tend to show even more contamination by comparison to measured concentrations on the Lincoln Park properties).

Recently, National Uranium Resource Evaluation (NURE) Data were made available to the public on CD-ROM. This database covers the conterminous western United States, and includes data gathered since 1974. The NURE program encompasses the Hydrogeochemical and Stream Sediment Reconnaissance (HSSR) survey which began in 1976 and ended in 1980. Stream sediment, soil, and surface and groundwater samples were

the bulk of the sample types collected. Originally, the samples were analyzed primarily for uranium. In 1977, analyses for additional elements were added. A representative sample is included for thorium as Figure 6-1. Lincoln Park is located within the Pueblo USGS quadrangle in which 4,522 records exist for the following sample types: spring water, lake water, stream water, wet sediment (from groundwater, streams, and lakes), dry sediment, and unknown sediments. These data were not collected with the intention of using samples as free as possible from contamination; they are simply a collection of all available analyses regardless of intent or analysis technique. A map showing data collection sites and average concentrations for thorium in ppm is included (Figure 6-1). Nickel had a high percentage (59.5) of samples testing below the LLD, therefore the conservative average concentration derived by forcing LLD values to zero is heavily weighted and the true value is higher. Copper, cobalt, and uranium all were assayed in at least 1,655 samples of which at least 94 percent exceeded the LLD.

Summary of Contamination for All Plaintiffs

All property is in the southern portion of Canon City, Colorado within a neighborhood known as Lincoln Park. Comparison of soil analysis results to background levels (Ebens and Shacklette, 1982; Shacklette and Boerngen, 1984; Kabata-Pendias and Pendias, 1984; Adriano, 1986; Bowen, 1979; Bowen, 1966; Rose, 1979; Aubert and Pinta, 1977; EPA, 1982; NCRP, 1987; Lowder, 1964; Vinogradov, 1959; and Bear, 1964) indicates elevated levels of arsenic, cobalt, copper, molybdenum, nickel, radium, uranium, and/or selenium. Those elements have been identified as contaminants released solely or in part by the Cotter Uranium Mill.

Arsenic

A total of 148 soil samples were analyzed for arsenic concentrations for properties north of the Cotter Mill. One laboratory had a lower limit of detection (LLD) of 10 ppm. All sample concentrations less than the LLD were assigned a concentration of half of the LLD. The samples average 11.4 ppm higher than the average level (7.2 ppm) for the conterminous United States. Arsenic levels are 2.6 times the background overall. The range maximum is 122.8 ppm above the background.

Cobalt

The Cotter Mill has been identified as the sole source of cobalt contamination in this area (Remedial Investigation). Results from the analysis of 61 soil samples provide an average of 26.9 ppm above the background of 9.1 ppm. The maximum concentration is 54 ppm, and the average multiplication factor over the background is 4.0. The average NURE cobalt concentration in the Pueblo Quadrangle is 12.06 ppm. The elevation in the plaintiffs' concentration over reference and NURE levels is strong evidence for the existence of cobalt contamination.

Copper

Copper concentrations were also determined for 61 soil samples. The range maximum is 91 ppm greater than the background (25 ppm), and the average is an increase of 3.9 ppm over the background with a multiplication factor of 1.16. The NURE level is an average of 30.42 ppm. This is an indication of contamination with specific areas of high concentration.

Molybdenum

Two analysis sensitivities were reported by the laboratories: 10 ppm and 1 ppm. A value of half of the corresponding LLD was applied when either was reported in the set of 156 samples. The net ppm (5.7)

and multiplication factor (6.9) averages are driven by a few very high values, e.g. the maximum of 121 ppm compared to a background of 0.97 ppm. Some samples contain levels that are far greater than the background, and the Cotter Mill has been identified as the sole source of this contaminant. Kabata-Pendias and Pendias (1984) report an average of 0.88 ppm (range of 0.4 - 2.8) in alluvial surface soils of the United States and an average of 2 ppm (range of 0.8 - 3.3 ppm) in various United States surface soils. They also report an average of 5.8 ppm (range of 1.5 - 17.8 ppm) in alluvial soils from areas of the western states showing molybdenum toxicity to grazing animals. That range maximum is greatly exceeded at some of the sampling locations.

Nickel

The Cotter Mill is the sole source of nickel contamination in the Lincoln Park area. None of the 61 samples has a concentration below the LLD. The average net concentration is 9.7 ppm above the background of 19 ppm. The range maximum is 52 ppm (or 33 ppm above background). The soil concentration exceeds the conterminous United States level by an average multiplication factor of 1.51 with a median of 1.42 and mode of 1.37. The data indicate a consistent and elevated level of nickel contamination.

Radium

Radium concentrations were measured as radium-226 in pCi/g, and translated to ppm to maintain consistency in this report. Radium and uranium data are shown in Tables 6-3, 6-4, and 6-5. The background of radium-226 is $0.66 \times 10^{-6}$ ppm based on the uranium-238 background of 1.8 ppm. The average radium-226 concentration on the affected properties is $0.93 \times 10^{-6}$ ppm, and the maximum concentration is $2.63 \times 10^{-6}$ ppm. In an undisturbed natural state, the activities of $^{238}U$ and $^{226}Ra$ should be equivalent, i.e. in a condition of secular equilibrium. For the tested

radionuclides, radium activity concentrations surpass those of uranium. This is evidence of a disturbance.

Selenium

A sum of 83 soil samples were analyzed for selenium. The maximum is 6.31 ppm above the background of 0.39 ppm. Measurements reported at the LLD were again assigned half of that value. The average selenium concentration for all properties is 0.5 ppm.

Uranium

Some uranium levels are particularly high, e.g. 20 and 33 ppm in sediments on the non-plaintiff property compared to a reference value of 1.8 ppm (i.e. factors of 11.1 and 18.3 above the background). Another notable sample from playground soil at a school has 6 ppm uranium. Sample statistical outputs for uranium and a radium to uranium comparison are given in Table 6-6.

Summary

The weight of the evidence for the Lincoln Park area supports the conclusion that arsenic, cobalt, copper, molybdenum, nickel, selenium, radium, and uranium are present in Lincoln Park soils at concentrations indicative of industrial contamination. Note that groundwater and surface water contamination are not considered in this study, however higher levels in stream sediments and irrigated land parcels suggests the importance of these pathways. Groundwater studies may be better indicators of uranium and radium contamination than the soil study. It is concluded that the natural condition has been disturbed in the uranium decay series. Samples taken in some of the plaintiffs' households (mainly attic dust) indicated contaminant infiltration of homes. Uranium concentrations peaked at 14 ppm, and radium-226 was measured as high as 12 pCi/g. Other contaminants found in those samples

were arsenic (82.4 ppm maximum), molybdenum (66 ppm maximum), and selenium (1.5 ppm maximum).

Table 6-3 Radium and Uranium Analyses for Houses 1-10

| Hazen Research, Inc. soil analyses for 10/93 and results | | | | | | | |
|---|---|---|---|---|---|---|---|
| House ID | Depth (inches) | U-238 ppm | Net ppm | U-238 pCi/g | Ra-226 pCi/g | U-Ra pCi/g | Net Ra pCi/g |
| 1 | 0 - 2 | 3 | 0.3 | 1.0 | 1.3 | -0.29 | 0.40 |
| 1 | 2 - 4 | 2 | -0.7 | 0.7 | 1.8 | -1.13 | 0.90 |
| 2 | 0 - 2 | 2 | -0.7 | 0.7 | 1.4 | -0.73 | 0.49 |
| 2 | 2 - 4 | 2 | -0.7 | 0.7 | 1.6 | -0.93 | 0.70 |
| 3 | 0 - 2 | 3 | 0.3 | 1.0 | 1.2 | -0.19 | 0.29 |
| 3 | 2 - 4 | 2 | -0.7 | 0.7 | 1.1 | -0.43 | 0.20 |
| 4 | 0 - 2 | 2 | -0.7 | 0.7 | 0.8 | -0.13 | -0.10 |
| 4 | 2 - 4 | 2 | -0.7 | 0.7 | 1.7 | -1.03 | 0.80 |
| 5 | 0 - 2 | 3 | 0.3 | 1.0 | 1.6 | -0.59 | 0.70 |
| 5 | 2 - 4 | 2 | -0.7 | 0.7 | 1.2 | -0.53 | 0.29 |
| 6 | 0 - 2 dup | 4 | 1.3 | 1.3 | 1.1 | 0.24 | 0.20 |
| 6 | 0 - 2 | 2 | -0.7 | 0.7 | 0.6 | 0.07 | -0.31 |
| 6 | 2 - 4 | 3 | 0.3 | 1.0 | 1.6 | -0.59 | 0.70 |
| 7 | 0 - 2 | 2 | -0.7 | 0.7 | 1.6 | -0.93 | 0.70 |
| 7 | 2 - 4 | 3 | 0.3 | 1.0 | 0.8 | 0.21 | -0.10 |
| 8 | 0 - 2 | 2 | -0.7 | 0.7 | 0.4 | 0.27 | -0.51 |
| 8 | 2 - 4 | 2 | -0.7 | 0.7 | 0.3 | 0.37 | -0.61 |
| 9 | 0 - 2 | 3 | 0.3 | 1.0 | 0.2 | 0.81 | -0.71 |
| 9 | 2 - 4 | 6 | 3.3 | 2.0 | 1 | 1.01 | 0.09 |
| 10 | 0 - 2 | 3 | 0.3 | 1.0 | 1.5 | -0.49 | 0.60 |
| 10 | 2 - 4 | 2 | -0.7 | 0.7 | 1.4 | -0.73 | 0.49 |
| Background arithmetic mean (ppm) | | | | | 2.7 | | |

Source: Shacklette and Boerngen, 1984

Table 6-4 Radium and Uranium Analyses for Houses 11-19

| Hazen Research, Inc. soil analyses for 10/93 and results | | | | | | | |
|---|---|---|---|---|---|---|---|
| House ID | Depth (inches) | U-238 ppm | Net ppm | U-238 pCi/g | Ra-226 pCi/g | U-Ra pCi/g | Net Ra pCi/g |
| 11 | 0 - 2 | 3 | 0.3 | 1.005516 | 1.4 | -0.394484 | 0.495 |
| 11 | 2 - 4 | 3 | 0.3 | 1.005516 | 1.4 | -0.394484 | 0.495 |
| 12 | 0 - 2 | 2 | -0.7 | 0.670344 | 1.5 | -0.829656 | 0.595 |
| 12 | 2 - 4 | 2 | -0.7 | 0.670344 | 1.3 | -0.629656 | 0.395 |
| 13 | 0 - 2 | 2 | -0.7 | 0.670344 | 2.1 | -1.429656 | 1.195 |
| 13 | 2 - 4 | 3 | 0.3 | 1.005516 | 0.7 | 0.305516 | -0.205 |
| 14 | 0 - 2 | 4 | 1.3 | 1.340688 | 0.6 | 0.740688 | -0.305 |
| 14 | 2 - 4 | 3 | 0.3 | 1.005516 | 1.3 | -0.294484 | 0.395 |
| 15 | 0 - 2 | 3 | 0.3 | 1.005516 | 1.3 | -0.294484 | 0.395 |
| 15 | 2 - 4 | 3 | 0.3 | 1.005516 | 0.6 | 0.405516 | -0.305 |
| 16 | 0 - 2 | 2 | -0.7 | 0.670344 | 1 | -0.329656 | 0.095 |
| 16 | 2 - 4 | 2 | -0.7 | 0.670344 | 1.2 | -0.529656 | 0.295 |
| 17 | 0 - 2 dup | 2 | -0.7 | 0.670344 | 1.3 | -0.629656 | 0.395 |
| 17 | 0 - 2 | 2 | -0.7 | 0.670344 | 1.3 | -0.629656 | 0.395 |
| 17 | 2 - 4 | 2 | -0.7 | 0.670344 | 2.6 | -1.929656 | 1.695 |
| 18 | 0 - 2 | 2 | -0.7 | 0.670344 | 0.8 | -0.129656 | -0.105 |
| 18 | 2 - 4 | 2 | -0.7 | 0.670344 | 2.5 | -1.829656 | 1.595 |
| 19 | 0 - 2 | 2 | -0.7 | 0.670344 | 0.2 | 0.470344 | -0.705 |
| 19 | 2 - 4 | 2 | -0.7 | 0.670344 | 0.4 | 0.270344 | -0.505 |
| Background arithmetic mean (ppm) | | | | | 2.7 | | |

Source: Shacklette and Boerngen, 1984

Table 6-5 Radium and Uranium Analyses for Houses 20-29

| Hazen Research, Inc. soil analyses for 10/93 and results | | | | | | | |
|---|---|---|---|---|---|---|---|
| House ID | Depth (inches) | U-238 ppm | Net ppm | U-238 pCi/g | Ra-226 pCi/g | U-Ra pCi/g | Net Ra-226 pCi/g |
| 20 | 0 - 2 | 2 | -0.7 | 0.670344 | 1.1 | -0.429656 | 0.195 |
| 20 | 2 - 4 | 1 | -1.7 | 0.335172 | 0.3 | 0.035172 | -0.605 |
| 21 | 0 - 2 | 1 | -1.7 | 0.335172 | 0 | 0.335172 | -0.905 |
| 21 | 2 - 4 | <1 | -2.7 | 0 | 0.2 | -0.2 | -0.705 |
| 22 | 0 - 2 | 1 | -1.7 | 0.335172 | 0 | 0.335172 | -0.905 |
| 22 | 2 - 4 | 2 | -0.7 | 0.670344 | 0 | 0.670344 | -0.905 |
| 23 | 0 - 2 | 1 | -1.7 | 0.335172 | 0 | 0.335172 | -0.905 |
| 23 | 2 - 4 | 1 | -1.7 | 0.335172 | 0 | 0.335172 | -0.905 |
| 24 | 0 - 2 | 2 | -0.7 | 0.670344 | 0 | 0.670344 | -0.905 |
| 24 | 2 - 4 | 7 | 4.3 | 2.346204 | 0.2 | 2.146204 | -0.705 |
| 25 | 0 - 2 | 1 | -1.7 | 0.335172 | 1.4 | -1.064828 | 0.495 |
| 25 | 2 - 4 | 2 | -0.7 | 0.670344 | 0 | 0.670344 | -0.905 |
| 25 | 2 - 4 dup | 1 | -1.7 | 0.335172 | 0.3 | 0.035172 | -0.605 |
| 26 | 0 - 2 | 2 | -0.7 | 0.670344 | 0.5 | 0.170344 | -0.405 |
| 26 | 2 - 4 | 2 | -0.7 | 0.670344 | 0.9 | -0.229656 | -0.005 |
| 27 | 0 - 2 | 1 | -1.7 | 0.335172 | 0.8 | -0.464828 | -0.105 |
| 27 | 2 - 4 | 1 | -1.7 | 0.335172 | 0.2 | 0.135172 | -0.705 |
| 28 | 0 - 2 | 2 | -0.7 | 0.670344 | 1.1 | -0.429656 | 0.195 |
| 28 | 2 - 4 | 2 | -0.7 | 0.670344 | 0.4 | 0.270344 | -0.505 |
| 29 | 0 - 2 | 2 | -0.7 | 0.7 | 0.8 | -0.13 | -0.10 |
| 29 | 2 - 4 | 8 | 5.3 | 2.7 | 0 | 2.68 | -0.91 |

Background arithmetic mean (ppm) 2.7
Source: Shacklette and Boerngen, 1984

Table 6-6 Summary Statistics for U-Ra Equilibrium

| Net Ra-226 in pCi/g compared to background | | U-Ra in pCi/g |
|---|---|---|
| Mean | 0.011 | -0.131 |
| Standard Error | 0.083 | 0.099 |
| Median | 0.095 | -0.2 |
| Mode | -0.905 | 0.335 |
| Standard Deviation | 0.645 | 0.776 |
| Variance | 0.416 | 0.603 |
| Skewness | 0.318 | 0.800 |
| Range | 2.6 | 4.611 |
| Minimum | -0.905 | -1.930 |
| Maximum | 1.695 | 2.681 |
| Sum | 0.695 | -7.970 |
| Count | 61 | 61 |
| Confidence Level(0.950000) | 0.162 | 0.195 |

CHAPTER 7
APPLICATION OF NEW TECHNIQUES TO EMERGENCY PLANNING

## Overview

Often, the first task in emergency planning is to find the shortest route from a hazardous area to a safe one. A GIS is set up to do this easily by simply selecting line segments and querying. The imbedded algorithm can calculate the segment length from one indexed point to another in the base map coordinate system. The concept of distance has historically been associated with the "shortest line between two points" which is frequently insufficient in a decision-making environment. The standard measurement unit used in most GIS is the "grid space" implied by superimposing an imaginary uniform grid over a geographic study area. The distance between locations is computed as the number of intervening grid spaces. In evacuation modeling this gives a deceptively simplified view of the situation evacuees would encounter such as undulations in horizontal and vertical planes and barriers to smooth traffic flow.

In many applications, however, the shortest route between locations may not always be a straight line. Even if it is straight, the geographic length of that line may not always reflect a meaningful measure of distance. Distance in evacuation planning is best defined in terms of movement expressed as travel-time. Distance-modifying effects may be expressed cartographically as "barriers" located within the space in which the distance is measured. Note that this implies that distance

is the result of some sort of movement over that space and through those barriers. Absolute barriers are those that completely restrict movement implying an infinite distance as travel-time between the locations they separate, unless a path around the barrier is available.

Distance Estimation

Many procedures used by decision support systems optimize distance-based objective functions. Researchers have suspected that simple interpoint distances are not always appropriate. As a result, locational analysis software (e.g., Goodchild and Noronha, 1983) often allows alternatives to symmetrical interpoint distances (e.g., time, or asymmetrical distance matrices to model one way streets).

Varying interpoint distances by a constant also has been the subject of empirical research because of differences between straight-line and network distances. Nordbeck (1963), for example, found that a constant could be applied to straight-line distance to compensate for measurement problems. Love and Morris (1972; 1979) also examined several methods (e.g., Euclidean and spherical) for producing estimates, and compared each to "actual" distances obtained from road map distance matrices and government agencies. The estimates were also parametrically altered. Simple Euclidean distances performed poorly when compared to their estimates.

Kolesar (1979) provides a distance approximation for deriving estimated travel times based on distances within U.S. urban areas.

$$D_{ij} = 1.15 \cdot \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

where 1.15 is an empirical constant.

The size and form of parameters change among various studies, and Ginsburgh and Hansen (1974) assert that values ranging between Euclidean

distance, and Euclidean distance x (1+SQRT(2)) are acceptable. This variability accentuates uncertainty in producing distance estimates, leads to problems with interpretation of results, and can lead to an inability to replicate spatial modeling experiments.

The entire estimation process may be obviated if "actual" distances in the form of digital networks are available for analysis. Information of this kind is available in the form of USGS Digital Line Graphs (DLG). These DLG data are an ideal source of reference data, because they are encoded at a high level of precision and are topologically validated.

Analysis outcomes are influenced by the scale of the source material (map base) or distance estimation procedure used in locational models. Both the value of the objective function and the geometrical configuration of the allocation are altered when different, but commonly employed, distance estimation procedures are used in the analysis. An important effect centers on the extreme generalization of some data sources.

## St. Lucie Nuclear Power Plant Evacuation Estimates

In this case study, the GIS capabilities of distance estimation and allocation of spatially located attributes are used to calculate populations within user-constructed geographic zones, project future populations and measure roadway lengths. These features in conjunction with evacuation modeling software are used to develop evacuation times estimates for the power plant.

This study presents evacuation estimates for the area surrounding the St. Lucie Nuclear Power Plant. The evacuation estimates presented herein reflect estimated 1992 resident, transient, and special

populations for the area within 10 miles of the St. Lucie plant. This effort includes consideration of currently designated routes and all other major roadways for both normal and adverse weather. Adverse weather is defined as: rain conditions, roads open, side streets passable, driveways clear, roadway capacities reduced 20 percent, and free flow speeds reduced 20 percent. The evacuation zone is subdivided per instructions from FP&L staff. Baseline 1990, estimated, and projected populations are described in previous chapters. Other dynamic population components are detailed below.

Transient Population

The total peak 1992 transient population within 10 miles of the St. Lucie Plant is estimated at about 46,711. This total includes an estimate of the number of seasonal visitors and daily transients. As in much of Florida, this region experiences significant fluctuations in population as thousands come to the area for the winter season (generally from Christmas to Easter) or for a summer or winter vacation. Many attractions and events are held throughout the year which draw thousands of people. Although few in number, major industries and colleges draw many workers and students every day. Transient population resulting from transportation by road, rail, waterway, and by air is estimated by calculating the average daily passengers at locations where vehicles or passenger counts are made. The transient population estimates and projections for the 10-mile (Evacuation Preparation Zone) EPZ are current as of December of 1992.

Tourists and seasonal visitors

The St. Lucie area experiences a significant influx of transient visitors during the winter months, from about December to April. This seasonal population includes persons with residence in another state who

come to the area for several months out of the year. Because their stay can be a lengthy one, planning officials have found it difficult to distinguish between permanent and seasonal residents. Seasonal residents may also claim permanent residence to avoid state income taxes. Other seasonal visitors include persons who come into the area for a short vacation, staying in seasonal housing units, overnight tourist accommodations, recreational camping areas or with friends and relatives. One of the more important cities in the plume pathway evacuation zone is Port St. Lucie. In 1990, the city had less than 450 hotel/motel units, no recreational vehicle parks, and no migrant labor camps (City of Port St. Lucie, 1992). The 1980 Census indicated that Port St. Lucie's housing stock had a permanent occupancy rate of 81.7 percent. This occupancy rate reflects a moderate to high degree of transient residents. Since 1980, the city has experienced a visible increase in younger families with children and in year-round residents. It is estimated that the occupancy rate has increased to approximately 85 percent to 86 percent during the 1980's and can be expected to remain at that rate. The city also estimates that approximately 95 percent to 96 percent of its housing units are occupied during the peak season. Therefore, seasonal population projections for Port St. Lucie are based upon an increase of approximately 10 percent above permanent residents. This anticipates that the majority of the seasonal influx is due to out-of-state residents moving to existing housing units within the city. Sector SSE, with an estimated 13,918 visitors in 1992, accommodates the largest number of overnight visitors due to the large number of recreational vehicle lots and condominium units on Hutchinson Island. Of the total, 9,586 persons visit sector-segment SSE 5-10. In segment 4-5 of the same sector, 4,013 persons are expected to visit on a peak day. This figure is greater than six of the nine 5-10 mile segments,

and eight of the twelve sectors for the entire ten miles. The 4-5 mile annulus is the most populous by far for one mile increments to five miles with a peak daily total of 6,338 persons. For a look at other factors affecting transient populations see Table 7-1.

Table 7-1 Recreational Population Near the St. Lucie Plant

| VISITORS TO RECREATIONAL FACILITIES WITHIN 10 MILES OF THE ST. LUCIE PLANT | | | |
|---|---|---|---|
| Name | Sector | Estimated Parking Capacity | Peak 1992 Population |
| Fort Pierce Inlet State Park | NNW 5-10 | 350 | 1400 |
| South Jetty Park | NNW 5-10 | 35 | 140 |
| Porpoise Beach | NNW 5-10 | 10 | 40 |
| South Beach Boardwalk | NNW 5-10 | 94 | 376 |
| Gulfstream Beach | NNW 5-10 | 10 | 40 |
| Surfside Park | NNW 5-10 | 100 | 400 |
| Frederick Douglas Mem. Park | NNW 5-10 | 50 | 200 |
| Middle Cove Beach | NNW 3-4 | 16 | 64 |
| Blind Creek Beach | N 1-2 | 16 | 64 |
| Walton Rocks Beach | ESE 0-1 | 150 | 600 |
| Hermans Bay Beach | SE 1-2 | 16 | 64 |
| Normandy Beach | SE 1-2 | 14 | 56 |
| Waveland Beach | SSE 5-10 | 70 | 280 |
| Jensen Beach | SSE 5-10 | 140 | 560 |
| Bob Graham Beach | SSE 5-10 | 20 | 80 |
| Alex's Beach | SSE 5-10 | 20 | 80 |
| Tiger Shores | SSE 5-10 | 25 | 100 |
| Savannas State Preserve | NW 3-5<br>WNW 3-4 | 64 | 256 |

Population at major industrial facilities

St. Lucie County has the highest unemployment rate of all of the counties in the 50-mile radius. The unemployment rate has steadily risen from 8.8 percent in 1988 to the present 13.8 percent in St. Lucie County. Martin County has suffered the same trend with an increase from 5.6 percent in 1988 to 9.8 percent in 1991. St. Lucie County's major private sector employers are HCA Lawnwood Regional Medical Center (1,146), Florida Power & Light (905), Winn-Dixie (755), Publix (641),

HCA Medical Center of St. Lucie (475), Southern Bell (429), Club Med/Village Hotel - Sandpiper (425), Harbor Federal (309), Flowers Bakery (275), and Becker/Indian River Foods (210) (Florida Department of Commerce, 1991). St. Lucie County last updated its employment figures by Traffic Analysis Zone (TAZ) in 1988. These figures were used to approximate the employment in industrial, commercial, and service sectors within 10 miles of the plant. The total employment for 1988 is estimated to have been 26,735. The projected employment for 1995 within 10 miles of the plant is 32,082. The major employers in Martin County are Martin Memorial Hospital (1,700), Grumman Aerospace (1,036), Armellini Express Lines (300), Indian River Plantation (350), First National Bank (350), Stuart News (275), TCT Turbo Combustor Technology (258), Florida Power & Light (240), Ebasco (240), and Dickerson, Inc. (200).

Most of the major industries surrounding the St. Lucie Plant involve citrus growing, packing and processing, construction or marine equipment. Facilities with at least 50 employees were included in this population segment. Many of the areas's largest industries depend on growth, such as construction and real estate. Other large employers, e.g. schools, governments, utilities, and retail stores, depend indirectly on growth. The steadiest local producer is agriculture, which employs more than 10,000 Treasure Coast workers. Freezes in the early and late 1980's forced many citrus growers south into St. Lucie County. Flowers Baking Company and Tropicana Products are 2 of the largest employers identified within ten miles, with a total of 400 (a significant increase from the year 1991) and 300 employees, respectively. FPL itself dismissed about 60 employees in St. Lucie County, including the nuclear plant, but sees no more cutbacks in the near future. Table 7-2 shows an analysis of employment by TAZ.

Table 7-2 Employment Within 10 Miles of St. Lucie by TAZ

| ESTIMATED 1988 EMPLOYMENT WITHIN 10 MILES OF THE ST. LUCIE PLANT | | | | | | | |
|---|---|---|---|---|---|---|---|
| | FRACTION | EMPLOYMENT | | | ADJUSTED EMPLOYMENT | | |
| TAZ | INCLUDED | INDUST. | COMMERCE | SERVICE | INDUST. | COMMERCE | SERVICE |
| 1 | 0.7 | 5 | 2 | 2 | 3.5 | 1.4 | 1.4 |
| 2 | 1 | 0 | 3 | 33 | 0 | 3 | 33 |
| 3 | 1 | 3 | 100 | 459 | 3 | 100 | 459 |
| 7 | 1 | 2 | 73 | 62 | 2 | 73 | 62 |
| 8 | 1 | 26 | 60 | 88 | 26 | 60 | 88 |
| 9 | 1 | 25 | 120 | 255 | 25 | 120 | 255 |
| 10 | 1 | 41 | 22 | 50 | 41 | 22 | 50 |
| 11 | 1 | 29 | 80 | 751 | 29 | 80 | 751 |
| 12 | 1 | 15 | 33 | 333 | 15 | 33 | 333 |
| 13 | 1 | 27 | 55 | 573 | 27 | 55 | 573 |
| 14 | 1 | 70 | 46 | 2802 | 70 | 46 | 2802 |
| 15 | 1 | 34 | 121 | 427 | 34 | 121 | 427 |
| 16 | 1 | 0 | 8 | 14 | 0 | 8 | 14 |
| 17 | 1 | 27 | 940 | 252 | 27 | 940 | 252 |
| 18 | 1 | 38 | 609 | 212 | 38 | 609 | 212 |
| 19 | 1 | 57 | 25 | 79 | 57 | 25 | 79 |
| 20 | 1 | 60 | 6 | 27 | 60 | 6 | 27 |
| 21 | 1 | 88 | 15 | 639 | 88 | 15 | 639 |

Special populations

Special populations which consist of subsets of the resident population include public school, private school, and college students; hospital and hospice patients; and inmates at correctional facilities. Most, but not all, of these special population members will be counted in the resident population, however their locations during school hours, inpatient duration, and incarceration sentences will be different.

Public and private schools in the two counties of interest are given by number and enrollment based on the 1990-1991 school year as shown in the following table.

Table 7-3 School Enrollment by County

| County | No.Public | No.Students | No.Private | No.Students |
|---|---|---|---|---|
| Martin | 26 | 11,808 | 10 | 2,012 |
| St. Lucie | 28 | 22,224 | 10 | 2,033 |

School enrollment for St. Lucie County was assigned to sector-segments in radii of 0-5, 5-10, and 10-20 miles from the St. Lucie Plant. These data were extracted from 1988 TAZ attributes. Sector-segment NW 5-10 contains 14,150 students, which is the most populous. The only students within 5 miles of the plant are in SSW 0-5 with a total of 500. There are no schools within 10 miles of the plant in the NNW sector. Other sector-segments that are known to have schools showed no students according to 1988 TAZ data: these are SW 0-5, W 5-10, WNW 0-5, and WNW 5-10. St. Lucie County's school system presently (Fall - 1992) has about 23,500 students; an increase of about 1,000 over the previous year. Four new schools opened in the county in 1992 to accommodate the growing student population. For example, Forest Grove Middle School in Fort Pierce opened in August, 1992 to about 1,200 students. Martin County has about half as many students with almost 12,000 tallied in 1992.

The current enrollment for Indian River Community College (IRCC) in Ft. Pierce is 9,483 students. There are now campuses for IRCC and Barry University in St. Lucie West. Indian River Community College is currently based in Fort Pierce, with branch campuses in Stuart, Vero Beach, St. Lucie West, and Okeechobee. The college has an attendance of 48,000 persons per year. Barry University, formerly a women's college based in Miami Shores, opened a branch campus in St. Lucie West in January of 1992. Their first official year started in October, 1992. Table 7-4 shows an analysis of school enrollment by TAZ.

Table 7-4 1988 St. Lucie School Enrollment by TAZ

| ST. LUCIE COUNTY SCHOOL ENROLLMENT YEAR: 1988 | | | |
|---|---|---|---|
| SECTOR | 0-5 MILES | 5-10 | 10-20 |
| SSW | 500 | 3000 | 0 |
| SW | 0 | 500 | 500 |
| WSW | 0 | 1900 | 0 |
| W | 0 | 0 | 0 |
| WNW | 0 | 0 | 500 |
| NW | 0 | 14150 | 3500 |
| NNW | 0 | 0 | 1000 |
| TOTAL | 500 | 19550 | 5500 |

Local health service centers include Lakewood Park Medical Center, Fort Pierce Community Health Center, Lawnwood Regional Medical Center, Harbour Shores Hospital, HCA Medical Center of Port St. Lucie, Savannas Hospital, Martin Memorial MediCenter of Port St. Lucie, Emergi-Center of Jensen Beach, Martin Memorial Medical Center, Coastal Surgical Care, Martin Memorial MediCenter of Palm City, Martin Memorial Hospital South, Martin Memorial MediCenter of Hobe Sound, Martin Memorial MediCenter of Indiantown, and Sandy Pines. Martin County has two main hospitals with the opening of Martin Memorial Hospital South on Salerno Road west of US 1 in September, 1992. This 100-bed facility is a satellite of the 236-bed Martin Memorial Medical Center in Stuart. HCA Medical Center of Port St. Lucie expanded with a 28-bed obstetrics unit in January, 1992. All of the counties combined have a total of eight hospices with only 48 beds. The following is an estimation of the number of inmates and patients residing in federal and state-operated institutions and considered nonresidents of the local area as of April 1, 1990.

Table 7-5 Special Populations by County

| County | City | Inmates/Patients |
|---|---|---|
| Martin | Stuart | 24 |
| | Unincorporated | 1,342 |
| St. Lucie | Unincorporated | 102 |

Transportation sources of transient population

The transient population resulting from the four basic modes of transportation is estimated by calculating the average daily number of passengers at locations on roads, waterways, rails, and airports where vehicles, vessels or passengers are counted.

Within 10 miles of the St. Lucie Plant, highways and roads are a major source of transient population. SR A1A, SR 707, and US 1 are major north-south arterials. SR A1A passes within approximately 1,000 feet of St. Lucie Plant on Hutchinson Island. SR 707 along the mainland coast is less than 2 miles from St. Lucie Plant at its nearest point. US 1 is not only a major arterial north and south, but also a focus of commercial activity in St. Lucie County. At its closest point, US 1 is approximately 4.8 miles from St. Lucie Plant.

At or near the 10-mile radius, 4 major river crossings concentrate traffic over the St. Lucie and Indian Rivers. These include the South Bridge, Jensen Beach Bridge, and Stuart Causeway from the mainland to Hutchinson Island and the Roosevelt Bridge on US 1 in Stuart. In February and March, traffic congestion in the Fort Pierce area and at the access points to Hutchinson Island is a severe problem; in fact, it is considered a limit to growth. Recommendations for an additional bridge crossing the Indian River have been made for the northern end of Hutchinson Island, within Fort Pierce City limits or in adjacent areas. As of December, 1992 this has not proceeded beyond preliminary investigation.

Traffic from the larger region comes within 10 miles of St. Lucie Plant on Florida's Turnpike, Route 91. At its closest point, the Florida Turnpike is approximately 7.5 miles from St. Lucie Plant. Traffic counts made in 1990 resulted in peak season Average Daily

Traffic (ADT) counts in each direction of 16,999 vehicles on the turnpike north of Okeechobee Boulevard, 17,683 vehicles south of Okeechobee Boulevard, and 21,168 vehicles south of Port St. Lucie Boulevard (St. Lucie County Metropolitan Planning Organization, 1991). The number of passengers for 1992 has been estimated and projected from these ADTs.

Peak season ADTs for 1990 were 29,443 vehicles per direction on I-95 north of Orange Avenue, 28,867 north of Okeechobee Road, 37,579 north of Midway Road, 34,994 north of St. Lucie West Boulevard, 26,851 north of Gatlin Boulevard, and 36,163 south of Gatlin Boulevard. I-95 has been completed in St. Lucie, Martin, and Palm Beach Counties and now spans the length of the state from the Georgia state line to south Miami.

Average daily traffic counts and passenger projections through 2030 are tabulated for 156 road segments in Martin County, and 216 road segments in St. Lucie County. Projected passenger populations can change as construction efforts allow increased capacities on congested roadways. Ongoing and anticipated projects for 1992 and 1993 are discussed below for St. Lucie and Martin Counties.

Projects for St. Lucie County include

1. widening the Prima Vista Boulevard bridge over the North Fork of the St. Lucie River,
2. widening Port St. Lucie Boulevard to six lanes west of the turnpike to Savage Boulevard,
3. widening Port St. Lucie Boulevard to four lanes between Savage and Darwin Boulevards,
4. improving Midport Road between US 1 and Port St. Lucie Boulevard, and
5. engineering design work on South 25th Street between Port St. Lucie and Fort Pierce, Midway Road between US 1 and South 25th Street, Jenkins Road and Lennard Road.

Prima Vista Boulevard in Port St. Lucie was widened to 5 lanes in 1992. The work scheduled for the bridge will widen the last 2-lane gap in the road which is one of Port St. Lucie's 2 major east-west routes. Another east-west route, the proposed Palmer Expressway through northern Port St. Lucie between US 1 and I-95, has been cancelled by the state Department of Transportation.

Projects for Martin County include

1. widening County Road A1A to 5 lanes from Jefferson Avenue to Indian Street,
2. paving Cove Road from State Road 76 to US 1 to improve travel between western Martin County and the Port Salerno area,
3. building Willoughby Boulevard from Monterey Road to Indian Street with developers of Willoughby Golf Club concurrently building the section from Indian Street to Salerno Road,
4. improving the US 1 - Jensen Beach Boulevard intersection near the Treasure Coast Square Mall, and
5. improving the intersection of Sewall's Point Road and East Ocean Boulevard.

The widening of County Road A1A from Indian Street to Monterey Road has been delayed by lack of money; but the intersection of County Road A1A, Monterey Road, and Palm Beach Road could be improved in 1993-1994.

State planners are studying another plan to improve travel between the east and west coasts by creating a 4-lane, high-speed highway along State Roads 710 and 70. State Road 710 links West Palm Beach, Indiantown, and Okeechobee. State Road 70, which links Fort Pierce and the Bradenton-Sarasota area, meets State Road 710 near Okeechobee.

The population associated with vessels navigating along the intracoastal waterway or rivers in the area, was estimated based on information obtained from the State Department of Transportation (DOT). The State DOT provided information on the number of bridge openings and number of vessels passing under the five bridges within the 10-mile

area. The number of yearly vessels was divided by 365 days to obtain a daily average. This number was then multiplied by 4 passengers per vessel to obtain the total population associated with waterways travel.

The Florida peninsula is transversed from Fort Myers to Stuart by the Okeechobee Waterway, a cross-land lock system providing access from the Gulf of Mexico to the Atlantic Ocean and Intracoastal Waterway. From September 1986 to September 1987, 9,639 vessels used the Waterway. The Waterway runs from Fort Myers Harbor to Lake Okeechobee. At Port Mayaca in Martin County, the Waterway enters the St. Lucie Canal. Heading north and east through Indiantown, the Canal connects with the South Fork of the St. Lucie River in Stuart. This eastern terminus of the Waterway lies within the 10-mile radius in sectors S and SSE. In 1987, the average daily number of passengers on ships going through the locks was estimated to be 106 persons.

Within 10 miles of St. Lucie Plant, the Florida East Coast Rail Line passes at a distance of 2 miles from the Plant. It carries no passengers. Although no airports exist within the 10-mile radius, both St. Lucie and Martin Counties have airports located between the 10- and 20-mile radii. The St. Lucie County Airport is located north of the City of Fort Pierce in sector NW. The Stuart/Martin County Airport is located south of Stuart in annular sector S 10-20. Its use is limited primarily to test flights for Grumman Aerospace.

Methodology: Transient Population

The transient population within 10 miles of the plant was estimated based on the number of seasonal overnight visitors and daily visitors. Overnight visitors include seasonal residents, and persons on vacation staying at hotels and motels, campgrounds, or with friends. Daily visitors may include those persons attending special events,

visiting major attractions, working in the area, or attending major colleges. In order to use such data in a GIS, the data should be accumulated by established geographic areas, e.g., census block groups, or digitized geographic areas, e.g., TAZ. The data must then be entered into a database spreadsheet with locational fields that are matched identically to the geographic file. A sample of projected peak daily seasonal visitors is included as Table 7-6.

Table 7-6 Transient Population Within 10 Miles of St. Lucie

| YEAR 2030 | PEAK DAILY TOURISTS AND SEASONAL VISITORS WITHIN 10 MILES OF THE ST. LUCIE PLANT | | | | | | |
|---|---|---|---|---|---|---|---|
| | DISTANCE (MILES) | | | | | | |
| DIRECTION | 0-1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-10 | 0-10 |
| N | 0 | 180 | 0 | 0 | 0 | 0 | 180 |
| NNE | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NE | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ENE | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| E | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ESE | 782 | 0 | 0 | 0 | 0 | 0 | 782 |
| SE | 0 | 1,161 | 0 | 0 | 0 | 0 | 1,161 |
| SSE | 0 | 0 | 0 | 717 | 8,993 | 16,933 | 26,643 |
| S | 0 | 0 | 0 | 36 | 311 | 5,533 | 5,880 |
| SSW | 0 | 0 | 9 | 851 | 804 | 7,511 | 9,175 |
| SW | 0 | 1 | 2 | 112 | 1,945 | 6,113 | 8,172 |
| WSW | 0 | 8 | 7 | 20 | 500 | 3,494 | 4,028 |
| W | 0 | 3 | 4 | 441 | 520 | 4,161 | 5,129 |
| WNW | 0 | 0 | 10 | 462 | 825 | 3,243 | 4,540 |
| NW | 0 | 0 | 0 | 190 | 303 | 18,093 | 18,585 |
| NNW | 0 | 0 | 0 | 64 | 0 | 11,181 | 11,245 |
| Total by Annulus | 782 | 1,353 | 31 | 2,892 | 14,201 | 76,261 | 95,520 |

Overnight population

The number of seasonal visitors staying at hotels and motels within 10 miles of the plant was calculated based on the number of units at each facility and the specific location of them. The total number of units was multiplied by an average occupancy rate of 2.0 persons per

room to calculate the total population associated with these overnight accommodations. Sources used to identify these tourist accommodations included telephone directories (Southern Bell, 1992), county planning documents, personal communication, and field surveys. Although St. Lucie County is not presently considered to be a primary tourist destination, indications are that the profile of the seasonal visitor to this community is changing rapidly, partially as a result of the introduction of the New York Mets major league baseball training complex, as well as the expanded tourist use of the area's aquatic resources (St. Lucie County Board of County Commissioners and Department of Community Development, 1990).

Local planning officials have found it difficult to accurately estimate the number of seasonal residents in the area, since many of them are residents for several months at a time. For purposes of estimating this population segment, a portion of all multi-family housing units was designated as seasonal, since this type of residential facility is most commonly associated with seasonal use. Interviews with management personnel of many of these developments indicated that they were unsure as to how many units were used only seasonally. Responses ranged from units being used essentially by all permanent residents, even on Hutchinson Island, to about 25 percent being used year-round.

Based on the various responses received, it was assumed that on Hutchinson Island about 33 percent of the multi-family developments were used strictly seasonally. Since the island, being on the ocean, probably accommodates the highest number of seasonal residents, the percentage of seasonal residences that was used on the island was not considered to be appropriate for estimating seasonal residents on the mainland. Seasonal visitors to Fort Pierce are concentrated along the northern tip of South Hutchinson Island. Since the mainland portion of

Fort Pierce is not expected to attract a significant number of seasonal visitors, only 5 percent of the units were estimated to be used seasonally. Other mainland areas south of Fort Pierce probably attract more seasonal visitors than fort Pierce itself, but still not as much as Hutchinson Island. Seasonal residents along this area of the mainland may be more concentrated along other inland water bodies, such as the St. Lucie and Indian Rivers. As a comparison between the number of visitors attracted to Hutchinson Island and those attracted to Fort Pierce's mainland, it was assumed that 20 percent of multi-family units in this area are used seasonally. The location and number of multi-family units were determined based on field surveys and Traffic Analysis Zone (TAZ) data. An occupancy factor of 2.0 persons per unit was used to estimate the seasonal resident population.

In order to estimate the number of seasonal visitors staying with friends and relatives, information on tourism in Florida was obtained (Fernald and Purdum, 1992). These data showed what percentage of visitors stay in the various types of overnight accommodations which are available. By comparing the percentage of visitors which stay with friends or relatives to the estimated number of seasonal visitors which use other types of accommodations, the total number of visitors staying with friends was estimated. This figure was compared to the 1990 resident population to calculate a ratio between those numbers. The resulting ratio of was applied to the resident population to disaggregate this seasonal category to each annular sector within the 10-mile area. Data used in the analysis are summarized in Table 7-7 which gives indirect indicators of transient populations.

Table 7-7 Indirect Transient Population Data by County

| County | Tot. licensed lodgings (units) | Apt. buildings (units) | Rooming houses (units) |
|---|---|---|---|
| Martin | 2,798 | 1,158 | 38 |
| St. Lucie | 4,813 | 1,272 | 249 |
| | Rental condominiums (units) | Transient apartments (units) | Occupancy rate (Percent) |
| Martin | 30 | 362 | 53 |
| St. Lucie | 909 | 289 | 73 |

Transient population at major employment facilities

The largest employers in the 10-mile area have been identified along with the number of employees at these facilities. It is reasonable to assume that many of these employees are probably also residents of the area. For this reason, it was assumed that about half of the employees live beyond the plant's 10-mile radius and would therefore contribute to the transient population segment.

Transient population at major colleges

The number of students attending colleges within 10 miles of the plant was obtained by survey. Since students attending the IRCC Fort Pierce campus may travel some distance, it was assumed that, as with employees, of the students attending college in the area, 50 percent of them live beyond the 10-mile area. Therefore, those students contribute to the total transient population estimate.

Methodology for Estimating Transient Population from Transportation

Transient population generated by transportation is comprised of 4 basic modes: highways, railroads, waterways, and airports. Traffic volume numbers are given by average daily total number of passengers for highway, rail, waterway, and air traffic.

Between 0 and 10 miles of St. Lucie Plant, travelers on major roads were estimated from the average daily traffic count (ADT) (Martin County Administrative Center Engineering Department, 1992; St. Lucie County Metropolitan Planning Organization, 1991) from 1990 at the sampling stations closest to the 10-mile radius (preferably at or just inside the 10-mile line). Major roads include interstate highways and state roads. Where ADT counts separated traffic by direction of flow, travel into the 10-mile radius has been used. Where the directions were combined, the ADT count was divided in half, on the assumption that traffic is evenly distributed in both directions. Numbers of vehicles were increased annually by incremental county growth rates to the 1992 estimate.

Estimates of transient population were derived from the number of vessels and drawbridge openings counted in 1987 on the Indian River and the St. Lucie River. The annual number of vessels recorded by bridge tenders was divided by 365 to reach a daily average. The number of vessels was then multiplied by 4 passengers per vessel to arrive at an average daily number of passengers for each bridge. These estimates do not include passengers on small craft which can pass beneath the drawbridges.

On the St. Lucie Canal, transient population has been derived from the lock master's records of annual total number of vessels. The annual total number has been divided by 365 days in the year, and multiplied by 4 persons per vessel to reach an average daily number of passengers. On the St. Lucie Canal, all vessels are counted as they pass through the locks.

Methodology: Evacuation Times Estimates

The Regional Evacuation Modeling System (REMS) is a computerized decision support system developed by a research team at the University of Florida Industrial and Systems Engineering Department for modeling the emergency evacuation of geographical areas due to natural, nuclear, or chemical emergencies. The software is a menu-driven set of programs specifically developed for 286 (IBM) and higher personal computer platforms running under DOS with enhanced graphics capabilities for displaying area evacuation road maps and providing time-lapsed evolution of the evacuation process. REMS is coded entirely in Turbo PASCAL.

The principal features of REMS are

1. inclusion of the time dimension into the evacuation process,
2. ability to quickly test many "what if" scenarios,
3. provision of evacuation patterns for different weather conditions,
4. generation of road usage information and indication of bottleneck highway segments,
5. generation of necessary statistics on the movement of regional traffic in time,
6. provision of sector analysis of traffic through simulation with visual display and data output,
7. cartographic display of the evacuation process dynamically in time,
8. calculation of summary statistics of the evacuation process intersection and road utilizations, first and last time of utilization, criticality in the evacuation process, capacity utilization, and
9. generation of the total evacuation time of each zone and the entire network.

Table 7-8 Necessary Databases Required by REMS

| INTERSECTION DATABASE |
|---|
| number |
| x, y coordinate |
| dynamic capacity (veh/hr) |
| ROAD DATABASE |
| name |
| beginning node |
| ending node |
| dynamic capacity (veh/hr) |
| length (miles) |
| static capacity (veh/mile) |
| average speed (mph) |
| POPULATION DATABASE |
| origin node number |
| number of vehicles to evacuate |
| ORIGIN-DESTINATION DATABASE (OPTIONAL) |
| origin node number |
| destination node number |
| number of vehicles to evacuate |
| RESPONSE CURVE DATABASE |
| beginning time |
| cumulative percent of population available for evacuation |

The user can interactively alter any of these databases in whole, or alter part of the data contained within the current model. Consequently, testing scenarios for different population values or changing transportation road network is very easy with REMS. The response curve database allows the user to define the evacuees' response behavior during an evacuation, and to redefine and retest that response.

Once the alterations are finalized, it only takes a few menu selections to run a new scenario and collect statistics on the current

experiment. REMS uses several solution procedures for insuring the validity of the evacuation time estimates. Some of these procedures are network flow theory, dynamic network flows, linear programming, and simulation. One final specialty of REMS that makes it a powerful decision support tool is that it will display the evacuation process on the evacuation transport network as it evolves in time. The user has the ability to freeze the graphical display at any time of the evacuation, study the graphics and statistics, and continue the progress of evacuation with a keystroke. Extensive statistics are generated and stored for report generations by REMS. A sample of translation of street names to link (arc or roadway) number designations used as REMS input is available in Table 7-9.

Results

Both normal (fair weather) and adverse weather conditions were analyzed. Adverse weather is characterized by rainy conditions with passable roadways. For this study, the effect on vehicle speed and dynamic roadway capacity was conservatively assumed to be a 20 percent reduction in both categories. Table 7-10 is analogous to Figure R-22 of the Emergency Plan, and sector coverages are designated to match Figure R-1 of that document. By direct comparison to data published in the previous update of that report, it is evident the total number of automobiles estimated for the entire 10-mile zone has increased by 98 percent; however the total evacuation time under normal conditions only increases by 16 percent. Although roadway improvements have been made in the interim, the use of a more sophisticated computer model may have contributed to apparent increase in evacuation efficiency. Among the eight tri-sector groupings, updated evacuation estimates range from

nearly 45 percent faster than the previous time to 25.5 percent slower (in a grouping with well over twice as many cars).

Table 7-9 Sample Key to Link (Arc) Designations

| Road Name and Link (Arc) Number for Each Street in the Evacuation Study | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 605 | (Oleander Blvd) | 32 | 54 | 74 | 81 | | | | | |
| 607A | (Angle Blvd) | 100 | 101 | 103 | | | | | | |
| 608 | (Lucie Blvd) | 106 | 120 | | | | | | | |
| 611B | (Edwards) | 60 | 61 | 70 | 71 | 72 | | | | |
| 615 | (Hawley) | 56 | 63 | 66 | 85 | 107 | | | | |
| 68 | (Orange Blvd) | 86 | 87 | 88 | 89 | 93 | 97 | 143 | | |
| 70 | (Okeechobee) | 67 | 69 | 83 | 84 | 91 | 92 | 122 | 506 | |
| 707 | (Dixie Hwy) | 10 | 11 | 13 | 14 | | | | | |
| 707 | (Indian River Drive) | 117 | 51 | 50 | 29 | 21 | 28 | 20 | 9 | |
| 707A | (Causeway Blvd) | 18 | 19 | | | | | | | |
| 709 | (Glades Cutoff) | 58 | 121 | | | | | | | |
| 712 | (Midway) | 44 | 43 | 42 | 41 | 40 | 30 | | | |
| 713 | (Kings Blvd) | 90 | 94 | 102 | 105 | | | | | |
| 716 | (Port St. Lucie Blvd) | 24 | 142 | 141 | 127 | 140 | 132 | | | |
| 723 | (Savannah Road) | 12 | 158 | | | | | | | |
| 732 | (Jensen Beach Causeway) | 8 | | | | | | | | |
| 778 | (Prima Vista Drive) | 35 | 47 | 505 | | | | | | |
| A1A | (Ocean Blvd) | 3 | 4 | 6 | 159 | 160 | | | | |
| Atlantic Beach Blvd | | 110 | 111 | 157 | | | | | | |
| Bayshore | | 34 | 145 | 37 | 46 | 36 | 125 | 502 | | |
| Becker | | 131 | 133 | 139 | | | | | | |

Figures 7-1 and 7-2 are maps providing a sequentially closer and more detailed view of the evacuation zone, terminating with a "skeletal" display of source areas, arcs, and nodes (intersections or destinations). Figure 7-3 is the evacuation network (EPZ) used by state and local planners (Figure R-1 of the Emergency Plan).

Table 7-10 St. Lucie Evacuation Time Estimates

| COUNTY AFFECTED | APPLICABLE SECTOR | ESTIMATED NUMBER OF AUTOMOBILES | POPULATION EVACUATION TIME ESTIMATES (0-10 Miles) | | | |
|---|---|---|---|---|---|---|
| | | | Normal Weather | | Adverse Weather | |
| | | | Min. | Hours | Min. | Hours |
| St.Lucie (SL) | R,Q,P | 46,486 | 464 | 7.73 | 530 | 8.83 |
| SL | Q,P,N | 50,053 | 385 | 6.42 | 444 | 7.40 |
| SL | P,N,M | 34,798 | 227 | 3.78 | 282 | 4.70 |
| SL | N,M,L | 42,815 | 409 | 6.82 | 513 | 8.55 |
| SL/Martin | M,L,K | 50,572 | 409 | 6.82 | 520 | 8.67 |
| SL/Martin | L,K,J | 47,386 | 386 | 6.43 | 478 | 7.97 |
| SL/Martin | K,J,H | 47,209 | 514 | 8.57 | 641 | 10.68 |
| SL/Martin | J,H,G | 29,712 | 359 | 5.98 | 443 | 7.38 |
| SL/Martin | R,Q,P,N,M,L, K,J,H,G | 136,510 | 552 | 9.20 | 683 | 11.38 |

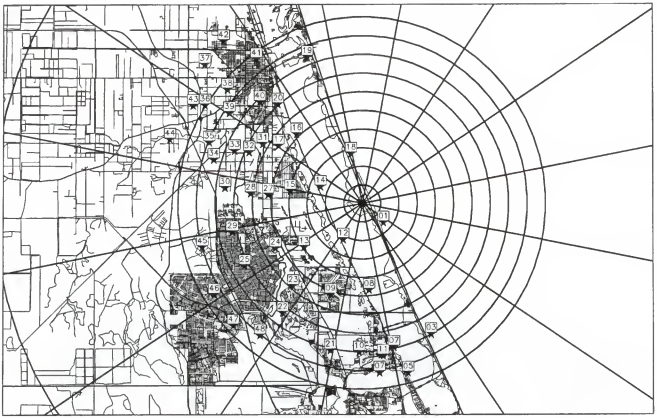


Figure 7-1 St. Lucie Evacuation Zone Showing Sources, Arcs and Nodes

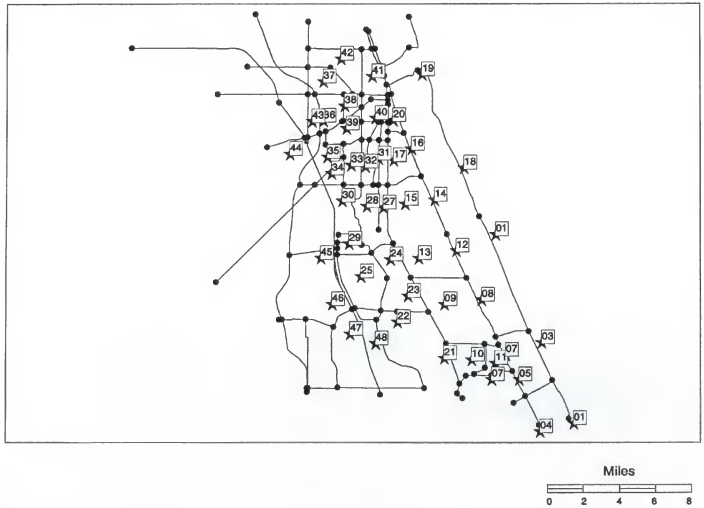


Figure 7-2 St. Lucie Evacuation Network Showing Arcs and Nodes

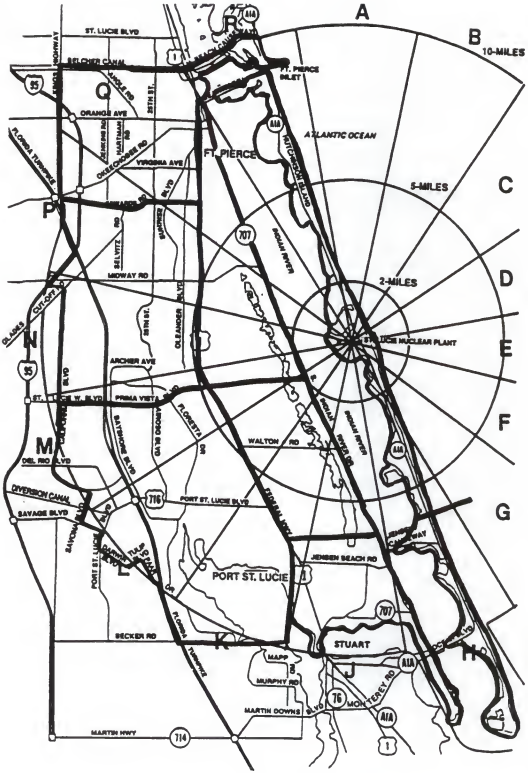


Figure 7-3 St. Lucie Emergency Planning Zones

# CHAPTER 8
# CONCLUSIONS AND RECOMMENDATIONS

## Summary Statement

The field of health physics draws upon other specialties for information and technologies relevant to emergency planning and radiation dose assessment. GIS technology has been used to produce highly detailed population, population density, housing unit, and housing unit density distributions. In addition, population projections and reconstructions can be generated by applying predictive equations to baseline data generated by area-weighting in the GIS. Population distributions around nuclear facilities are used as input data to atmospheric dispersion models for estimating offsite radiation doses due to routine and accidental releases of radioactive materials. Detailed case studies demonstrated the power of GIS for spatial analysis of not only population data, but also environmental samples, and demographic data useful for evacuation modeling. Populations are reported as point estimates with no regard for error in GIS manipulation or input data. The problem stems from the resistance of software vendors to incorporate uncertainty accounting algorithms within their programs, and reluctance to release proprietary codes to outside researchers. It is important for health physicists to realize that they can effectively perform many of their own GIS analyses, that the use of census levels that are too aggregated may drastically affect decisions based upon GIS output, and

that errors and uncertainties are generally ignored in the GIS industry and must be semi-quantitatively estimated.

The crucial point on hardware is to buy a system with the fastest processor available and a large hard drive with fast access and data transfer times. Multiple processor systems are only necessary for extremely large geographical areas, large attribute files, and frequent analysis. For most purposes, an off-the-shelf personal computer of high quality will suffice. Select a brand name model since some vendors cut costs by using lower grade components in models that seem equivalent, but will perform poorly. Get the fastest processor available coupled with a large (gigabytes in capacity) hard drive with a fast access time and rapid transfer rate. Another necessity is a high speed CD-ROM drive to access Bureau of the Census data.

## Recommended Procedures

### Population Baseline Distributions

Use the FIPS county codes to find the TIGER files for counties in the region of interest. Download the files from CD-ROM and disable the "read only" code (attrib -r *.*). The CD can be purchased from the Census Bureau, or borrowed from a federal repository (preferred to save cost). Use the Import/Export module to modify geographic files for ATLAS GIS. The street maps are useful for visualization and address matching. A boundary translation program must be used to extract census boundary information. BdryBNA is one such program that produces census polygons in the BNA format. Import the BNA files into the GIS and merge with the county boundaries included in the software. Translated geographic files and some attribute files are available from vendors,

but usually at a great cost. It is recommended that the relatively inexpensive translation software be purchased instead.

Now the population data must be incorporated. Obtain the appropriate State CD from the same source as above. Use the extraction program supplied on the disk to download a database file containing only the desired population data. The population data file must contain at least one field identical to the BNA file so that they can be matched. The geographic file and its matched attribute file can be queried using standard polygon on polygon overlay procedures. The specific commands will vary depending upon the GIS software. Construction of a sector-segment grid is done by establishing an origin data point (usually latitude/longitude coordinates for the center of the plant), specifying buffer rings at the chosen radial distances, and laying down buffer lines to form sectors. Each sector-segment must be a separate buffered region that can be queried for summary statistics.

Projections

It is recommended that the non-demographer obtain data estimates and projections professionally prepared by a State agency such as Florida's Bureau of Economic and Business Research. These documents are inexpensive and updated periodically. Use the county-level populations for projection years to generate growth ratios. Create new fields in the population database for each projection year and apply the growth ratio for each county to its census blocks in the spreadsheet. Next, construct the sector-segment grid as before and query the selected regions for the new area-weighted projections. Each sector-segment projection will be a composite of those projected blocks from the various counties that comprise it. If a researcher has an abundance of demographic data (especially migration data) and does not wish to use

the simple procedure designed for this study (above), then the techniques described in Chapter 3 should be addressed. However, once the projections are made for counties, the GIS procedure remains the same. The solution to the problem of saturation lies in whether or not one accepts the following assumptions. Assume that the highest population density for a census block in the base census year (e.g., 1990) is a maximum that the local population cannot or does not wish to exceed. Next, assume that adjacent or nearby high-density blocks are the more desirable to attract an excess populace. In the database write a script appropriate for the software of choice that calculates the projected block subtracts the population that exceeds the density maximum, and forces that excess to the second highest density block. Saturated blocks must be forced to stagnancy.

A second problem is that of null blocks. One can observe from the case study projections that regions with zero as a baseline can never grow using growth equations. In otherwise highly populated areas, such blocks usually involve water, farmland, industrial areas, etc., but this is not always the case. If the GIS user has good information about planned developments, the null regions can be "seeded" with families to generate future populations.

Reconstructions

Population reconstructions are essentially intercensus population estimations. Over such periods (decades) the method of choice among arithmetic, geometric, and exponential rates of change becomes inconsequential. More sophisticated methodologies suffer the fate of ageing data, i.e., boundaries of regions change (e.g., census tracts and zip codes), data reverts from electronic to hard copy to hard-to-get, location names change, and so on. GIS begins to lose its appeal when

large amounts of data must be entered into databases by hand, and obsolete regions must be digitized into a geographic file. The problem with population reconstructions around old nuclear weapons plants is that those populations are sometimes very small. A researcher in dire need of very accurate data might well be served by hiring a staff to gather the data by hand from administrative records.

Error Handling

A standard 50-mile sector-segment grid used for offsite dose assessment consists of fourteen annular segments (0-1, 1-2, 2-3, 3-4, 4-5, 5-6, 6-7, 7-8, 8-9, 9-10, 10-20, 20-30, 30-40, and 40-50 miles) and sixteen radial sectors centered on the major compass points for a total of 224 sector-segments. A typical carefully constructed grid was found to have an arithmetic mean of 0.13 percent error per sector-segment. The population of sector-segment errors is normally distributed with a standard deviation of 0.037. This means that a researcher interested in a sector-segment in the 40-50 mile annulus may include or exclude about 0.23 square miles of geography on average, and 68 percent of the sector-segments will range from 0.16 to 0.3 square miles in error. This may be significant depending upon the spatial distribution and density of the local population. The estimated population will have errors associated with 1) the census count itself, 2) the encoded census boundaries, 3) sector-segment grid construction, and 4) area-weighting of the population. Of these, categories 1 and 2 are beyond a user's control. The census count error is used in error propagation as reported by the Bureau of the Census for the decennial census of interest. The encoded census boundary error is assumed to be negligible. Category 3 is one that the user can attempt to deal with, and error due to category 4 can be minimized by use of the smallest possible census boundaries, e.g.,

the block level for the 1990 census, in the analysis. An estimate of the population error attributed to area-weighting can be attained by selecting all census boundaries intersected by the boundary lines of the sector-segment of interest. These are the area-weighted regions, and their population assignments can theoretically approach 100 percent error. A conservative treatment is to assign 100 percent error to each, sum for all intersected regions, and divide by the sector-segment total population for the fractional error. A more reasonable approach would be to assign 50 percent error in place of the 100 percent above and proceed as stated.

LIST OF REFERENCES

Adriano, D. C. Trace Elements in the Terrestrial Environment. New York, NY; Springer-Verlag; 1986.

Antenucci, J. C. Maryland Automated Geographic System: Expansions and Applications in State Government. In Urban, Regional, and State Applications, Cambridge, MA; Harvard Library of Computer Graphics 1979 Mapping Collection; Laboratory for Computer Graphics and Spatial Analysis; 1979.

Armstrong, M. P. Distance imprecision and error in spatial decision support systems. In International Geographical Information Systems (IGIS) Symposium: The Research Agenda; Proceedings; Volume II. Aangeenbrug, R. T.; Schiffman, Y. M. (eds). Arlington, VA; NASA, 1987.

Aubert, H. and Pinta, M. Trace Elements in Soils. New York, NY; Elsevier Scientific Publishing Company; 1977.

Bear, F. E. Chemistry of the Soil: 2nd Edition. London; Chapman and Hall; 1964.

Blakemore, M. Generalization and error in spatial data bases. Proc. Auto-Carto 6:313-322; 1983.

Bowen, H. J. M. Trace Elements in Biochemistry. New York, NY; Academic Press; 1966.

Bowen, H. J. M. Environmental Chemistry of the Elements. New York, NY; Academic Press; 1979.

Bowie, S. H. U.; Thornton, I. (eds). Environmental Geochemistry and Health: Report to the Royal Society's British National Committee for Problems of the Environment. Dordrecht, The Netherlands; D. Reidel Publishing Company; 1985.

Brauner, A. Population Estimates: Nuclear Power Plant Nearby Population Concentrations. Washington, DC; U.S. Nuclear Regulatory Commission; 1982.

Brunsdon, C. F.; Carver, S. The accuracy of digital representations of 2D and 3D geographical objects: a study by simulation. In GIS, Spatial Modelling and Policy Evaluation. Fischer, M.; Nijkamp, P. (eds). Heidelberg; Springer-Verlag; 1993.

Brunsdon, C. F.; Carver, S.; Openshaw, S.; Charlton, M. A review of methods for handling error propagation in GIS. Proc. 1st European GIS Conf.(EGIS'90); Amsterdam; 1990.

Brunsdon, C. F.; Openshaw, S. Simulating the effects of error in GIS. Chapter 7 in Geographical Information Handling--Research and Applications. Mather, P. M. (ed). Chichester,UK; John Wiley & Sons, Ltd.; 1993.

Bureau of Economic and Business Research (BEBR); Stanley K. Smith and Ravi Bayya. Projections of Florida population by county 1990-2020. College of Business Administration, University of Florida. BEBR 24:(2) Bulletin No.96; July, 1991.

Burrough, P. A. Principles of GIS for Land Resource Assessment. Oxford; Clarendon Press; 1986.

Buttenfield, B. Treatment of the cartographic line. Cartographica 22(2):1-26; 1985.

Callen, M.; James, I.; Mason, D. C.; Quarmby, N. A test-bed for experiments on hierarchical data models in integrated geographic information systems. Proc. Workshop Spatial Data Processing using Tesseral Methods; NERC Unit for Thematic Information Systems; University of Reading; Reading, UK; 1986.

Carver, S. J.; Brunsdon, C. F. Practical error handling in GIS map overlay. Proc. 3rd European GIS Conf.(EGIS'91); Brussels; 1992.

Carver, S.; Openshaw, S. Handling error propagation in GIS: A proposal for enhancing the Arc/Info tool-kit. NERRL Research Report. CURDS, Newcastle University; Newcastle, UK, 1990.

Casillas, P. Data aggregation and the P-median problem in continuous space. Discussion Paper No. 35. Iowa City, IA; Department of Geography, University of Iowa; Iowa City, IA; 1983.

Census of Population and Housing, 1990. Public Law (P.L.) 94-171 Data (Florida) machine-readable data files prepared by the Bureau of the Census. Washington; The Bureau (producer and distributor); 1991a.

Census of Population and Housing, 1990. Public Law (P.L.) 94-171 Data (Tennessee) machine-readable data files prepared by the Bureau of the Census. Washington; The Bureau (producer and distributor); 1991b.

Census of Population and Housing, 1990. Public Law (P.L.) 94-171 Data Technical Documentation prepared by the Bureau of the Census. Washington; The Bureau (producer and distributor); 1991c.

Census of Population and Housing, 1990. Summary Tape File 1 on CD-ROM Technical Documentation prepared by the Bureau of the Census. Washington; The Bureau (producer and distributor); 1991d.

Census of Population and Housing, 1990. Summary Tape File 1 on CD-ROM (Florida) machine-readable data files prepared by the Bureau of the Census. Washington; The Bureau (producer and distributor); 1991e.

Census of Population and Housing, 1990. Summary Tape File 1 on CD-ROM (Tennessee) machine-readable data files prepared by the Bureau of the Census. Washington; The Bureau (producer and distributor); 1991f.

Chappell, W. R.; Petersen, K. K. Molybdenum in the Environment (in two volumes). New York, NY; Marcel Dekker; 1976.

Chan, K.; White, D. Map Algebra: An object oriented implementation. In International Geographical Information Systems (IGIS) Symposium: The Research Agenda; Proceedings; Volume II. Aangeenbrug, R. T.; Schiffman, Y. M. (eds). Arlington, VA; NASA, 1987.

Chen, A. The Geographics Information Systems Laboratory. Berkeley, CA; Lawrence Berkeley Laboratory, Center for Building Science News; Spring:7, 1995.

Chrisman, N. R. A theory of cartographic error and its measurement in digital data bases. Proceedings: Auto-Carta 5. Falls Church, VA; American Society of Photogrammetry and American Congress on Surveying and Mapping; 1982.

Chrisman, N. R. On storage of coordinates in geographic information systems. Geo-Processing 2:259-270; 1984.

City of Port St. Lucie. Comprehensive Plan. St. Lucie County; 1992.

Deuker, K. J. Geographic data encoding issues; Technical Report No.43. Iowa City, IA; University of Iowa Institute of Urban and Regional Research; 1975.

Douglas, D.; Peucker, T. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Can. Cartogr. 10:112-122; 1973.

Ebens, R. J. and Shacklette, H. T. Geochemistry of Some Rocks, Mine Spoils, Stream Sediments, Soils, Plants, and Waters in the Western Energy Region of the Conterminous United States: Statistical Studies in Field Geochemistry. Geological Survey Professional Paper 1237. Washington, DC; United States Government Printing Office; 1982.

Engineering Department; Martin County Administrative Center. ADT Raw Data for 1992. Martin County, FL; 1992.

Environmental Protection Agency (EPA). Final Environmental Impact Statement for Remedial Action Standards for Inactive Uranium Processing Sites (40 CFR 192); Washington, DC; EPA; 1982.

Environmental Systems Research Institute, Inc. (ESRI). Understanding GIS: The ARC/INFO® Method. Harlow; Longman; 1993.

Farmer, D. G.; Rycroft, M. J. Computer Modelling in the Environmental Sciences. Oxford; Oxford University Press; 1991.

Fernald, E. D.; Purdum, E. D. Atlas of Florida. Tallahassee, FL; Institute of Science and Public Affairs, Florida State University; 1992.

Gahegan, M. N. An efficient use of quadtrees in a geological information system. Int. J. Geogr. Inf. Sys. 3(3):201-214; 1989.

Gardiner, V. Stream networks and digital cartography. Cartographica 19(2):38-44; 1982.

Geographic Reference File--Names, 1990. (Census Version) machine-readable data file prepared by the Bureau of the Census. Washington; The Bureau (producer and distributor); 1991.

Geotrans, Inc.; Rocky Mountain Consultants, Inc.; Eri Logan, Inc. Remedial Investigation: Cotter Corporation Uranium Mill Site. Prepared for The State of Colorado Department of Law: Office of the Attorney General; February, 1986.

Ginsburgh, V.; Hansen, P. Procedures for the reduction of errors in road network data. Operational Research Quarterly 25:321-322; 1974.

Gittings, B. M.; Sloan, T. M.; Healey, R. G.; Dowers, S.; Waugh, T. C. Meeting expectations: a review of GIS performance issues. Chapter 6 in Geographical Information Handling - Research and Applications. Mather, P. M. (ed). Chichester,UK; John Wiley & Sons, Ltd.; 1993.

Goodchild, M. F. The aggregation problem in location-allocation. Geographical Analysis 11:240-255; 1979.

Goodchild, M. F. The effects of generalisation in geographical data encoding. In Map Data Processing. Freeman, H.; Pieroni, G. (eds). New York, NY; Academic Press; 1980.

Goodchild, M. F. Towards an enumeration and classification of GIS functions. In International Geographical Information Systems (IGIS) Symposium: The Research Agenda; Proceedings; Volume II. Aangeenbrug, R. T.; Schiffman, Y. M. (eds). Arlington, VA; NASA, 1987.

Goodchild, M. F.; Gopal, S. (eds). The Accuracy of Spatial Databases. London; Taylor and Francis; 1989.

Goodchild, M. F.; Guoging, S. C.; Shiren, Y. Development and test of an error model for categorical data. Int. J. Geogr. Inf. Sys. 6:87-104; 1992.

Goodchild, M. F.; Noronha, V. Location-Allocation for Small Computers: Monograph 8. Iowa City, IA; Department of Geography, University of Iowa; 1983.

Goodchild, M. F.; Parks, B. O.; Steyaert, L. T. Geographic Information Systems and Environmental Modeling. Oxford; Oxford University Press; 1993.

Gough, L. P.; Shacklette, H. T.; Case, A. A. Element Concentrations Toxic to Plants, Animals, and Man: Geological Survey Bulletin 1466. Washington, DC; United States Government Printing Office; 1979.

Harding, T. J.; Hopkins, S. Polygon Overlay Concepts. Technical Report EPCC-CC93-10; Edinburgh Parallel Computing Centre and Department of Geography; University of Edinburgh; 1993.

Haygarth, P. M.; Jones, K. C. Atmospheric Deposition of Metals to Agricultural Surfaces. In Biogeochemistry of Trace Metals. D. C. Adriano (ed). Boca Raton, FL; Lewis Publishers; 1992.

Hendley, D. J. The scope for parallel computing in GIS. Mapping Awareness 4(5):7-10; 1990.

Hillsman, E. L.; Rhoda, R. Errors in measuring distances from populations to service centers. Annals of the Regional Science Association 12:74-88; 1978.

Hoffman, J. D.; Gunnells, G. B.; McNeal, J. M. United States Geological Survey Digital Data Series DDS-1, National Geochemical Data Base: National Uranium Resource Evaluation Data for the Conterminous Western United States; Electronic Data on CD-ROM; 1991.

João, E.; Herbert, G.; Openshaw, S.; Rhind, D. Magnitude and significance of generalisation and its effects. Proc. 3rd European GIS Conf. (EGIS'92); Munich; 1992.

João, E.; Herbert, G.; Rhind, D.; Openshaw, S.; Raper, J. Towards a generalization machine to minimize generalization effects within a GIS. Chapter 8 in Geographical Information Handling--Research and Applications. Mather, P. M. (ed). Chichester,UK; John Wiley & Sons, Ltd.; 1993.

Jackson, A. P.; Alloway, B. J. The Transfer of Cadmium from Agricultural Soils to the Human Food Chain. In Biogeochemistry of Trace Metals; Edited by D. C. Adriano, Boca Raton, FL; Lewis Publishers; 1992.

Jenks, G. F.; Lines, computers, and human frailties. Annals, Association of American Geographers 71(1); 1981.

Kabata-Pendias, A.; Pendias, H. Trace Elements in Soils and Plants. Boca Raton, FL; CRC Press; 1984.

Knill, J. Geographical Information Systems: the environmental view. Chapter 2 in Geographical Information Handling - Research and Applications. Mather, P. M. (ed). Chichester, UK; John Wiley & Sons, Ltd.; 1993.

Kolesar, P. J. Travel time and travel distance models: Chapter 6 in Fire Department Deployment Analysis. New York; North Holland; 1979.

LaGrega, M. D.; Buckingham, P. L.; Evans, J. C. Hazardous Waste Management. New York; McGraw-Hill; 1994.

Laurini, R.; Thompson, D. Fundamentals of Spatial Information Systems. The A.P.I.C. Series No. 37. San Diego, CA; Academic Press, Inc.; 1994.

Li, Z.; Openshaw, S. Algorithms for automated line generalisation based on a natural principle of objective generalisation. Int. J. Geogrf. Inf. Sys. 6:373-389; 1992.

Love, R. F.; Morris, J. G. Modelling inter-city road distances by mathematical functions. Operational Research Quarterly 23:67-71; 1972.

Love, R. F.; Morris, J. G. Mathematical models of road travel distances. Management Science 25(3); 1979.

Lowder, W. M.; Raft, P. D.; Beck, H. L. Field spectrometric investigations of environmental radiation in the U.S.A. In The Natural Radiation Environment. Adams, J. A. S.; Lowder, W. M. (eds). Chicago, IL; University of Chicago Press; 1964.

Luckey, T. D.; Venugopal, B. Metal Toxicity in Mammals (in two volumes). New York, NY; Plenum Press; 1977.

Maling, D. M. How long is a piece of string? The Cartographic Journal 5:147-156; 1968.

Mueller, P. Geographic Information System for radiologic emergency response. Frankfort, KY; Conference of Radiation Control Program Directors, Inc. (CRCPD) Newsbrief; January 2-3; 1995.

Müller, J. C. The concept of error in cartography. Cartographica 24:1-15; 1987.

Müller, J. C. Generalization of spatial databases. In Geographical Information Systems: Principles and Applications. Maguire, D. J.; Goodchild, M. F.; Rhind, D. W. (eds). Harlow; Longman; 1991.

Müller, J. C.; Zeshen, W. Area-patch generalisation: a competitive approach. Cartogr. J. 29:137-144; 1992.

National Council on Radiation Protection and Measurements (NCRP). Report No. 94: Exposure of the Population in the United States and Canada from Natural Background Radiation. Bethesda, MD; NCRP; 1987.

National Council on Radiation Protection and Measurements (NCRP). Report No. 118: Radiation Protection in the Mineral Extraction Industry. Bethesda, MD; NCRP; 1993.

Nordbeck, S. Computing distances in road nets. Papers of the Regional Science Association 12:207-220; 1963.

Nyerges, T. GIS Research Needs Identified During a Cartographic Standards Process: Spatial Data Exchange. In International Geographical Information Systems (IGIS) Symposium: The Research Agenda; Proceedings; Volume I. Aangeenbrug, R. T.; Schiffman, Y. M. (eds). Arlington, VA; NASA, 1987.

O'Conaill, M. A.; Mason, D. C.; Bell, S. B. M. Spatiotemporal GIS techniques for environmental modelling. Chapter 11 in Geographical Information Handling - Research and Applications. Mather, P. M. (ed). Chichester,UK; John Wiley & Sons, Ltd.; 1993.

Oheme, F. W., Ed. Toxicity of Heavy Metals in the Environment (in two parts). New York, NY; Marcel Dekker; 1978.

Olea, R. A. Optimum mapping techniques using regionalised variable theory: Series on Spatial Analysis, No. 2. Lawrence, KS; Kansas Geological Survey; 1975.

Openshaw, S. A methodology for using models for planning purposes. Environ. Plann. A. 11:879-896; 1979.

Openshaw, S. Learning to live with errors in spatial databases. In The Accuracy of Spatial Databases. Goodchild, M. F.; Gopal, S. (eds). London; Taylor and Francis; 1989.

Openshaw, S.; Charlton, M.; Carver, S. Error propagation: a Monte-Carlo simulation. In Handling geographic information: Methodology and Potential Applications. Masser, I.; Blakemore, M. (eds). London; Longman; 1991.

Perkal, J. On the length of empirical curves: Discussion paper No.10. Ann Arbor, MI; Michigan Inter-University Community of Mathematical Geographers; 1966.

Peuquet, D. J. A Conceptual Framework and Comparison of Spatial Data Models. Cartographica 21(4):66-113; 1984.

Pierzynski, G. M.; Sims, J. T.; Vance, G. F. Soils and Environmental Quality. Boca Raton, FL; Lewis Publishers (CRC Press); 1994.

Purves, D. Trace-Element Contamination of the Environment. New York, NY; Elsevier Scientific Publishing Company; 1977.

Raymondo, J. C. How to count illegals, state by state. American Demographics 10(9); 1988.

Raymondo, J. C. Population Estimation and Projection: Methods for Marketing, Demographic, and Planning Personnel. New York, NY; Quorum Books; 1992.

Rhind, D. W. Generalisation and realism within automated cartographic systems. Can. Cartogr. 10:51-62; 1973.

Rhind, D. W. GIS--A research agenda. Int. J. Geogr. Inf. Sys. 2(1):23-28; 1988.

Robinson, V. B.; Miller, R.; Klesh, L. Issues in the use of expert systems to manage uncertainty in geographic information systems. In International Geographical Information Systems (IGIS) Symposium: The Research Agenda; Proceedings; Volume II. Aangeenbrug, R. T.; Schiffman, Y. M. (eds). Arlington, VA; NASA, 1987.

Rose, A. W.; Hawkes, H. E.; and Webb, J. S. Geochemistry in Mineral Exploration: 2nd Ed. London; Academic Press; 1979.

Runyon, T. Integration of GIS and GPS technologies for environmental characterization and remediation of radiologically contaminated sites. Proceedings of the topical meeting: Environmental Radiation: How Clean is Clean? Springfield, IL; Prairie State Chapter of the Health Physics Society; 1994.

St. Lucie County Board of County Commissioners and Department of Community Development. St. Lucie County Comprehensive Plan Update, Future Land Use Element. St. Lucie County, FL; 1990.

St. Lucie County Metropolitan Planning Organization. Year 1990 Traffic Counts for St. Lucie County. St. Lucie County, FL; 1991.

Sathisan, S. K.; Chagari, A. K. Population estimation techniques for routing analysis. Las Vegas, NV; Proceedings of the 5th Annual International Conference on High-Level Radioactive Waste Management (Part 2 of 4):658-665; 1994.

Sejkora, K. J.; Most, M. L. Global Positioning System in Radiological Environmental Monitoring. Nuclear Plant Journal (March-April):48-53; 1993.

Severson, R. C.; Shacklette, H. T. Essential Elements and Soil Amendments for Plants: Sources and Use for Agriculture: USGS Circular 1017. Washington, DC; United States Government Printing Office; 1988.

Shacklette, H. T.; Boerngen, J. G. Element Concentrations in Soils and Other Surficial Materials of the Conterminous United States: Geological Survey Professional Paper 1270. Washington, DC; United States Government Printing Office; 1984.

Shyrock, H. S.; Siegel, J. S. The Methods and Materials of Demography. United States Bureau of the Census. Washington, DC; United States Government Printing Office; 1973.

Sinisgalli, A. 1980 Residential Population Estimates, 0-80 Kilometers for Nuclear Power Plants. Washington, DC; U.S. Nuclear Regulatory Commission; 1982.

Smythe-Wright, D. Ocean Circulation and Climate. Swindon; UK Natural Environment Research Council; 1991.

Steinitz, C.; Parker, P.; Jorden, L. Hand-drawn overlays: Their history and prospective uses. Landscape Architecture 66(8):444-455; 1976.

Stewart, J. C. Geographic criteria for the siting of low-level waste disposal sites. In International Geographical Information Systems (IGIS) Symposium: The Research Agenda; Proceedings; Volume II. Aangeenbrug, R. T.; Schiffman, Y. M. (eds). Arlington, VA; NASA, 1987.

TIGER/LINE (TM) Census Files, 1990. Machine-readable data file prepared by the Bureau of the Census. Washington; The Bureau (producer and distributor); 1991a.

TIGER/LINE (TM) Census Files, 1990. Technical Documentation prepared by the Bureau of the Census. Washington; The Bureau (producer and distributor); 1991b.

Tomlinson, R. F. Technology Alternatives and Technology Transfer. In Computer Assisted Cartography and Geographic Information Processing: Hope and Realism. D. Douglas; A. R. Boyle (eds). Ottawa, Canada; Canadian Cartographic Association and Department of Geography, University of Ottawa; 1982.

Tyrwhitt, J. Surveys in planning. In Town and Country Planning Textbook. Association for Planning and Regional Reconstruction (ed). London; The Architectural Press; 1950.

United States Bureau of the Census (USBC 1975). Population Estimates and Projections: Current Population Reports, 1975. Series P-25, No. 541. Washington, D.C.; U.S. Dept. of Commerce, Social and Economic Statistics Administration, Bureau of the Census; 1975.

United States Bureau of the Census. Current Population Reports; Series P-25; No. 984; Evaluation of Population Procedures for Counties: 1980. Washington, DC; United States Government Printing Office; 1986.

United States Bureau of the Census. Statistical Abstract of the United States: 1988; 108th edition. Washington, DC; United States Dept. of Commerce; 1987.

United States Bureau of the Census. Current Population Reports; Local Population Estimates; Series P-26; No. 85-TN-C; Estimates of the Population of Tennessee Counties and Metropolitan Areas: July 1, 1981, to 1985. Washington, DC; United States Government Printing Office; 1988.

United States Bureau of the Census. Current Population Reports; Series P-25; No. 1039; Patterns of Metropolitan Area and County Population Growth: 1980 to 1987. Washington, DC; United States Government Printing Office; 1989.

United States Bureau of the Census. Current Population Reports; Series P-23; No. 166; Perspectives on Migration Analysis. Washington, DC; United States Government Printing Office; 1990a.

United States Bureau of the Census. Current Population Reports; Series P-25; No. 1053; Projections of the Population of States, by Age, Sex, and Race: 1989 to 2010. Washington, DC; Government Printing Office; 1990b.

United States Bureau of the Census. Current Population Reports; Series P-25; No. 1063; State and Local Agencies Preparing Population and Housing Estimates. Washington, DC; United States Government Printing Office; 1990c.

United States Geological Survey. Canon City, Colorado 1:100,000-scale metric topographic map: 30 x 60 Minute Quadrangle; 1982.

United States Geological Survey. Pueblo, Colorado 1:250,000-scale metric topographic map: 1 x 2 Degree Quadrangle; 1989.

Vandegrift, G. F.; Reed, D. T.; Tasker, I. R. Environmental Remediation: Removing Organic and Metal Ion Pollutants. Washington, DC; American Chemical Society; 1992.

Vinogradov, A. P. Geochemistry of Rare and Dispersed Elements in Soils. London; Chapman and Hall; 1959.

Visvalingam, M.; Whyatt, J. The Douglas-Peucker algorithm for line simplification: Re-evaluation through visualization. Comput. Graphics Forum 9:213-228; 1990.

Waldron, H. A. Metals in the Environment. New York, NY; Academic Press; 1980.

White, M. S. Technical requirements and standards for a multipurpose geographic data system. The American Cartographer 11(1):15-26; 1984.

Worboys, M. F. A model for spatio-temporal information. Proc. 5th Int. Symp. Spatial Data Handling; Charleston, SC; 1992.

Young, J. M. Measuring the quality of geographic databases. In International Geographical Information Systems (IGIS) Symposium: The Research Agenda; Proceedings; Volume I. Aangeenbrug, R. T.; Schiffman, Y. M. (eds). Arlington, VA; NASA, 1987.

## BIOGRAPHICAL SKETCH

The author is a baccalaureate graduate of the University of Florida Department of Zoology. He subsequently participated in genetics research for the United States Department of Agriculture at the Insects Affecting Man and Animals Research Laboratory in Gainesville, Florida. Research duties included the induction of chromosomal aberrations such as inversions and reciprocal translocations using ionizing radiation, and the handling of radioisotope tracers. He received the degree of Master of Science in 1990 as a graduate student in the Department of Environmental Engineering Sciences at the University of Florida. He has been a doctoral student in health physics from that time to present. He has twice taught a graduate level course in radiation dosimetry, and lectured on health physics and health physics measurements. During graduate study he has pursued professional consulting projects involving radiation dose estimation, pathway analysis, decontamination and decommissioning of retired facilities, population estimation and projection, emergency evacuation time estimates, and radioactive waste disposal. Clients include Florida Power Corporation, Florida Power & Light Company, CH2M Hill, OptoMechanik, and Cargill Fertilizer.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

W. Emmett Bolch, Chair
Professor of Environmental Engineering Sciences

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Paul Chadik
Assistant Professor of Environmental Engineering Sciences

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

David Hintenlang
Associate Professor of Nuclear Engineering Sciences

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William Properzio
Associate Professor of Environmental Engineering Sciences

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Charles Roessler
Professor of Environmental Engineering Sciences

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May, 1997

Winfred M. Phillips
Dean, College of Engineering

Karen A. Holbrook
Dean, Graduate School

LD
1780
1997
.B619